ESSAYS ON ASSOCIATIONS OF BUILT ENVIRONMENT WITH USE OF RESTAURANTS AND FOOD STORES AND FOOD PURCHASE IN THE UNITED STATES

Ke Peng

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of City and Regional Planning in the College of Arts and Sciences.

Chapel Hill
2018

Approved by:

Nikhil Kaza

Penny Gordon-Larsen

Daniel A. Rodríguez

Catherine Zimmer

T. William Lester

Danielle Spurlock

ABSTRACT

Ke Peng: Essays on Associations of Built Environment with Use of Restaurants and Food Stores and Food Purchase in the United States
(Under the direction of Nikhil Kaza)

Planners often focus on increasing the availability of supermarkets to improve diet-related behaviors. It is, however, unclear how people use different types of outlets and how household purchase food with different outlet options, both of which in turn affect diet-related behavior and ultimately health outcomes. The limited number of previous studies have reported weak or null associations between food availability and use of food outlets or food purchase. These inconsistencies may be due to a lack of concern on the broader built environment in which food outlets situated that account for unobserved heterogeneity in the convenience/attractiveness of using such outlets by individuals.

To address the limitation, I used individual-level frequency of use of restaurants and food stores data from the Coronary Artery Risk Development in Young Adults, household-level food purchase data from the Nielsen Homescan Consumer Panel, with geographically-linked food outlet locations, the other built environment factors, and neighborhood sociodemographic characteristics. I sought to explicitly quantify cross-sectional associations between neighborhood food availability, the broader built environment context, and the use of restaurants and food stores and food purchase.

I found that individuals living in neighborhoods with more sit-down restaurants frequented sit-down restaurants more and greater degree of neighborhood street connectivity was

associated with less frequent use of neighborhood fast food restaurants and grocery stores. I found that households which lived in neighborhoods with greater neighborhood street connectivity reported more expenditures on fresh fruits and vegetables; households with greater numbers of convenience stores in their neighborhood purchased less fruits. I also found that sit down restaurants were more likely to be situated in inner city neighborhoods compared to other neighborhoods in 2011 than previous observational years, using the Twin Cities Region of Minnesota between 1993 and 2011 as a case study.

My results suggest that it is important for future studies and policy interventions to account for the broader built environment context of the food environment. Interventions focusing on increasing the availability of supermarkets or decreasing the availability of fast food restaurants in order to increase access to healthy food should proceed with caution.

ACKNOWLEDGEMENTS

I feel immense gratitude to many people, but I must first mention the one person who essentially enabled me to complete my doctoral program. My adviser, Dr. Nikhil Kaza, saved me from that dreaded “ABD” (“All but dissertation”) status by rescuing me during the most painful time of my doctoral program and gave me what I needed, when I needed it, including patience, reassurance, trust, and thoughtful feedback. I would have left the program without his substantial support. He treated me as an equal and a peer, which made me feel like I truly accomplished something worthwhile in my dissertation and with my time in North Carolina.

I also owe a deep debt to Dr. Penny Gordon-Larsen. As a faculty member of the Department of Nutrition at UNC, she gave me much more than what she should have. She spent tons of time on guiding me how to write my research proposal and how to do scientific writing. I’m not sure whether words can adequately express my gratitude for her help.

I am also indebted to my other committee members, including: Daniel A. Rodríguez, who walked me through the key steps in two of the three papers, and challenged me to think about the big picture and the implications of my work for improving current and future work in my chosen fields of research; Catherine Zimmer, who met with me on a regular basis and patiently walked me through difficult modeling approaches; T. William Lester, who gave me a lot of helpful ideas when I did not know how to move forward; and, Danielle Spurlock who provided rigorous feedback and encouraged me to think about how and where real people shop and eat.

This work was made possible by the Carolina Population Center and the Spatial

Analytic Unit, especially Marc Peterson, who developed and managed the geospatial database I used for my dissertation. I'm also grateful to my peers for their support, especially Gwen Kash, Lindsay Maurer-Braun, Yan Chen, Chaosu Li, and Allison Forbes, who provided the kind of friendship only other PhD students can give, as well as those in my "cohort cuddle", who created a safe space in which I could share my ideas and struggles.

I am also thankful for the unconditional support of my parents, parents in-law and my twin son and daughter. I appreciate their patience and understanding, and their pushing me to do my doctoral program because it will give me happiness.

Most of all, I am beyond fortunate to have the unwavering support and help of my husband, who has been my cheerleader since day one. Nothing I accomplished at UNC would have been possible without his unshakeable confidence in my abilities, and his saint-like patience for my moments of frustration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	American Community Survey
CARDIA	Coronary Artery Risk Development in Young Adults
CI	Confidence Interval
D&B	Dun & Bradstreet
GIS	Geographic Information System
MSA	Metropolitan Statistical Area
OR	Odds Ratio
SD	Standard Deviation
SE	Standard Error
SES	Socioeconomic Status
SIC	Standard Industrial Classification
SLD	Smart Location Database
U.S.	United States
UPC	Universal Product Code
VIF	Variation Inflation Factor

INTRODUCTION

Statement of the problem

Placing supermarkets in areas with limited access to healthy food (Cummins et al., 2014, Cummins et al., 2005, Sadler et al., 2013, Wrigley et al., 2003, Elbel et al., 2015, Dubowitz et al., 2015, Zenk et al., 2017) or implementing zoning restrictions on fast food restaurants (Sturm and Hattori, 2015) have been promoted as ways to improve diet and reduce the risk of obesity. Although planners are putting more emphasis on developing local solutions (e.g., farmers' markets, food pantries, innovative financing mechanisms) to solve the problems of social exclusion from healthy food or being disproportionally exposed to unhealthy food, their focus on retail-provision intervention on food consumption patterns is rare (Eating, 2008). More research is needed to inform practicing planners of the need to encourage food retail development that is appropriate in type, number, and scale to the neighborhood it serves. Evidence based policy interventions need to account for food purchase behaviors of different groups in the context of different choices conditioned by the built environment.

Although disadvantaged groups (e.g., low socioeconomic status and unemployed) may be more constrained to their neighborhood food retail outlets (Burgoine et al., 2014), some evidence suggested or implied that understanding the food shopping behavior should be extended to all. The studies on the ubiquity of energy-dense snack foods and beverages found unhealthy foods are widely available in a variety of retail stores including those whose primary business is not food such as gas stations and pharmacies (Popkin, 2011, Christian and Rashad, 2009), which may threat the diet quality of everyone. The studies on "choice architecture" suggested that altering the physical environment may alter people's food-related behavior in a predictable way without forbidding any options (e.g., restrict the number of fast food restaurant) or significantly changing economic incentive (e.g., beverage tax), which can influence the behavior of many people simultaneously and they are not targeted or tailored to specific individuals (Bucher et al., 2016). Because of the ability of planners to mobilize various spatial elements (e.g. travel routes, and modes, signage, parking restrictions, street frontage) at a broader scale, it is possible to improve the possibilities of people's access to healthy food by reshaping the built environment, rather than allowing people to passively adapt to the food environment with unhealthy foods.

Key literature review and conceptual framework

What are the mechanisms through which the number and type of food stores are associated with their use? In particular, does the context of the neighborhood built environment (e.g., street connectivity and regional destination accessibility) contribute to the explanation of healthy food use and dietary behaviors? These questions can be addressed by examining empirical evidence for the associations between neighborhood food availability and the frequency of use of restaurants and food stores, food purchasing, and measurement of neighborhood food availability under an ecological framework (to properly conceptualize the many food environments and conditions that influence food choices). This section elaborates my conceptual framework based on the review of the literature concerning built environment and food purchase.

Previous studies of the association between neighborhood food availability and use of food store and food purchase generated mixed results. Several have provided support for a positive association between neighborhood restaurant/food store availability and the frequency of use (Boone-Heinonen et al., 2011a, Forsyth et al., 2012, Laxy et al., 2015), whereas others have not (Richardson et al., 2011, Jeffery et al., 2006a, Oexle et al., 2015, Laska et al., 2010b).

Similarly, some studies found a strong or modest positive association between neighborhood supermarket availability and the purchase of healthy food (Volpe et al., 2013, Handbury et al., 2015, Lin et al., 2014), but others did not find such an association (Kyureghian et al., 2013).

One of the reasons which may have contributed to these mixed results is that previous studies gave little emphasis to the broader built environment context in which food outlets are situated (Kerr et al., 2012). The broader neighborhood built environment context encompasses a range of physical and social elements that constitute the structure of a neighborhood and which may influence how people use food outlets (Papas et al., 2007). For example, neighborhoods unsafe for walking or which lack transit access, can result in difficulty for individuals in accessing supermarkets and an overreliance on nearby convenience stores, which may in turn affect diet (Rose and Richards, 2004). The lack of empirical evidence reported on the food-related broader built environment context in urban area draws attention to the unfortunate fact that issues of food acquisition (and the related systems or processes) are less visible than other urban problems (e.g., housing or employment), thus food issues are frequently regarded as agricultural and rural issues (Pothukuchi and Kaufman, 1999) in the field of urban planning. With a better knowledge of the broad built environment context in which restaurants and food stores operate and how people interact with these food outlets in such a broader built environment, it is possible to reframe food decisions without coercing the individual and to design policies that are both effective and unobtrusive (Schwartz et al., 2017).

Effective interdisciplinary collaborations between different fields (e.g., urban planning, public health nutrition, physical activity research), the availability of secondary retail food outlet data (e.g., food inspection registries, InfoUSA, Yellow Pages) with address information, nationwide food purchase data collected by business information companies (e.g., Nielsen), and the access to Geographic Information System (GIS) techniques have created opportunities for planners to study the built environment influencing the use of restaurants and food stores, which is critical to understanding the relationship between the built environment and healthy eating.

The number of neighborhood food outlets in the United States increased substantially in the past decades regardless of type (Kim et al., 2015), which greatly facilitates access to foods and beverages (Gordon-Larsen, 2014). Recent studies have indicated that easy access to all food, rather than lack of access to specific healthy foods, may be a more important factor in explaining the increase in obesity (Ver Ploeg, 2010). This assertion suggested two necessities that I address in my dissertation. The first of these is accounting for the presence of the complementary outlet options when examining the use of one specific type of food outlet. The second is assessing the composition of food outlets in these neighborhoods, in addition to the number of each type of food outlet. Relative measures, such as the ratio or the proportion of various types of food retail outlets, may be as or more important to diet-related behaviors than the number of outlets because they offer neighborhood residents competing options (Clary et al., 2015, Mercille et al., 2012, Rummo et al., 2015).

As neighborhoods have become increasingly diverse in social composition and physical form over the past decades in the U.S., the use of restaurants and food stores in one type of neighborhood (e.g., urban) could be greatly different from the use of such outlets in another type of neighborhood (e.g., rural) as the distribution of types of stores in different neighborhoods may be different in regards to number, type, and distance. Identifying those neighborhoods experiencing great change in food outlet number and composition, e.g., those with consistently relatively fewer supermarkets or greater availability of fast food or sit-down restaurants, help us better make note of neighborhoods deserving special concern in the future.

Food store usage in different types of neighborhoods (e.g. urban vs suburban) might be different (Richardson et al., 2011) as the composition, types of stores, distance thresholds might be different. Furthermore, neighborhoods are becoming increasingly diverse in their social composition as well as physical form and it is important to tease out the relationships between urban change and change in dietary behaviors. Identifying those neighborhoods experiencing great change in food outlet composition, e.g., those with consistently relatively fewer supermarkets or greater availability of fast food or sit-down restaurants, help us better make note of neighborhoods deserving special concern in the future.

From the literature review, I summarize the relationships among built environment and

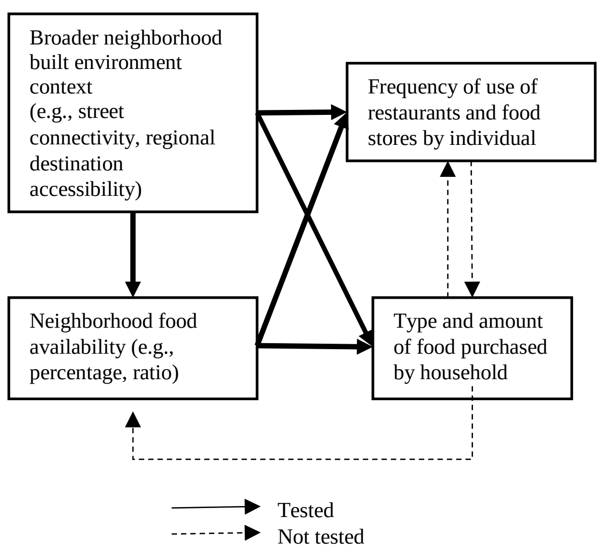

Figure I. Conceptual framework of dissertation

dietary behavior (see Figure I). Greater number of stores is associated with higher frequency of use of food outlets and greater amount of food purchased. However, other built environmental factors, such as street connectivity or regional destination accessibility are correlated with frequency of use of food outlets and the amount of food purchased. Over time, changes in relative compositions of stores could be associated with different neighborhood types. Taken together, examining this set of relationships has implications for how we can refashion the built environment to promote healthy eating habits.

As shown in Figure I, I hypothesize that greater neighborhood food availability is associated with higher frequency of use of food outlets and greater amount of food purchased. I further hypothesize that other built environmental factors, such as street connectivity or regional destination accessibility, are positively or associated with frequency of use of food outlets and the amount of food purchased. Third, I hypothesize neighborhoods with distinct built environmental and sociodemographic characteristics differed in the availability of restaurants and food stores and some neighborhoods experienced change in the availability of restaurants and food stores over time.

Research questions and hypotheses

- Research question 1: Does the neighborhood food availability relate to the use of restaurants and food stores by individuals and food purchasing of households?
    - Hypotheses to be investigated:
        - H1a: Individuals in neighborhoods with a greater number of fast food restaurants/sit-down restaurants/grocery stores use such restaurants or food stores more frequently.
        - H1b: Households in neighborhoods with greater numbers of neighborhood supermarkets will report greater expenditures on fresh fruits and vegetables; households in neighborhoods with greater numbers of neighborhood convenience stores will report less expenditures on fresh fruits and vegetables;

- Research question 2: Do other built environment characteristics, such as neighborhood street connectivity or regional destination accessibility, help to explain the use of restaurants and food stores by individuals and food purchasing by households?
    - Hypotheses to be investigated:
        - H2a: Neighborhood street connectivity moderates the relationship between the numbers of particular types of food outlets and the frequency of use of those neighborhood outlets.
        - H2b: Households in neighborhoods with greater neighborhood street connectivity, regional destination accessibility, neighborhood destination diversity and availability of neighborhood destinations will report greater expenditures on fresh fruits and vegetables.
- Research question 3: Do neighborhoods differ in the (relative) availability of restaurants and food stores? What type of neighborhoods experienced changes in the (relative) availability of neighborhood restaurants and food stores over time?
    - Hypotheses to be investigated:
        - H3a: Urban core and inner city neighborhoods consistently will have greater percentages of sit-down restaurants relative to total sit-down restaurants and fast food restaurants than other neighborhoods.
        - H3b: Suburban edge and high-income neighborhoods consistently will have greater percentages of supermarkets relative to total supermarkets, grocery stores and convenience stores than other neighborhoods.
        - H3c: The percentages of sit-down restaurants and supermarkets will increase over time regardless of neighborhood type.

Each of these research questions and hypotheses are examined in a separate chapter, with different datasets.

Paper 1: Home neighborhood food availability, street connectivity and frequency of use of neighborhood restaurants and food stores

In the first chapter of my dissertation, I examine the cross-sectional association of GIS-measured availability of neighborhood restaurants and food stores with the self-reported frequency of use of these food outlets, and whether such an association is modified by neighborhood street connectivity, using a large and diverse population-based cohort of middle-aged U.S. adults. Previous studies (Richardson et al., 2011, Forsyth et al., 2012, Boone-Heinonen et al., 2011a) have focused on the overall frequency of restaurant or food store use without measuring whether the restaurants/food stores that were used were within or outside of study participants' neighborhoods. Neighborhood food availability might be only weakly associated with the overall frequency of use if most participants used restaurants/food stores in other settings, such as school or workplace neighborhoods (Richardson et al., 2011, Burgoine and Monsivais, 2013). My work is the first to identify how neighborhood food outlets are used by residents using population-based data. My work is also in line with the previous work examining the trend towards fewer at-home family meals and increasing numbers of meals in eat-out places and a growing interest in understanding how people interact with neighborhood restaurants.

Previous studies have indicated that people in highly connected (or walkable) neighborhoods were more likely to participate in transport walking (e.g., walking to work or a grocery store) than were those in poorly connected neighborhoods (Witten et al., 2012, Ewing and Cervero, 2001), as connectivity affects the directness of travel and the number of alternative routes (Thornton et al., 2011). I conjectured that street connectivity, reflecting the convenience of accessing food resources, might modify the strength of the relationship between neighborhood restaurant/food store availability and frequency of using those food resources.

I found a positive association between the GIS-measured number of neighborhood sit-down restaurants and the self-reported frequency of using such restaurants. I also observed an inverse association between neighborhood street connectivity and the self-reported frequency of using neighborhood fast food restaurants. My work is the first to find (on a population level, through the use of a population-based dataset of medium-aged individuals) that the neighborhood built environment is related to the use of fast food restaurants and sit-down restaurants. My results indicate that built environment factors (e.g., number of sit-down restaurants and street network pattern) in the immediate home neighborhood were related to the use of neighborhood restaurants.

Paper 2: Built environment and the purchase of fruits and vegetables in United States households

The frequency of use of food outlets, however, does not indicate the type of food people ate at those restaurants or purchased in those stores. Food purchase is more directly linked to dietary behavior (e.g., food intake) than is the frequency of using outlets (since you can always buy a small amount of food such as a snack or a bottle of soda in neighborhood fast food restaurants). Although a couple of previous studies (Kerr et al., 2012, Laska et al., 2010a) suggested that food purchasing occurs in a broader built environment and people traveled sizeable distance for food, in the first chapter I only examine how residents used the restaurants and food stores in the immediate neighborhood and do not assess the use of restaurants and food stores outside of the home neighborhood. It is possible that the increasing number of different types of food outlets near a person’s residence over time may result in less food purchasing during their commute to or from work or during travel to other destinations. It is also possible that people obtain food from a variety of locations, many of which are outside of their local community.

To remedy this limitation, in the second chapter, I study the total amount of food purchased by a household regardless of the outlet location. In this study, I provide a first national examination of the cross-sectional associations between the built environment and expenditures on fresh fruits and vegetables in the continental U.S using Nielsen Homescan Consumer Panel Dataset for 2010. I focus on a range of built environmental factors that could facilitate or hinder food purchasing behaviors in retail stores (e.g., warehouse club stores, mass merchandisers and supercenters, grocery stores, convenience, drug, and dollar stores) in addition to the commonly examined availability of neighborhood supermarkets, such as availability of neighborhood convenience stores, regional destination accessibility, availability of neighborhood destinations, neighborhood destination diversity, and neighborhood street connectivity. I expected to obtain more information concerning where people purchase food by linking more neighborhood-level and region-level residential built environmental factors to food purchases.

I find that an increased number of neighborhood supermarkets was not associated with greater expenditures on fresh fruits and vegetables. My work provided evidence that interventions focusing exclusively on increasing the density of supermarkets in a neighborhood may not be as likely to induce better dietary behaviors through food purchasing. I found that households living in neighborhoods with greater numbers of neighborhood convenience stores reported less expenditures on fresh fruits. My results corroborated those of a recent study (Rummo et al., 2017) which also suggested that future initiatives to promote healthy food consumption behaviors should pay more attention to the relative availability of convenience stores than the relative unavailability of grocery stores and supermarkets . Policies to incentivize increase of healthy food offerings in the convenience stores is another solution to be considered. Additionally, regional destination accessibility was positively associated with expenditures on fresh fruits, which implied that households tended to use food stores outside their home neighborhoods to purchase fresh fruits because food opportunities were linked to other places encountered in weekly or daily routine travels beyond the home neighborhood (Kerr et al., 2012, Clifton, 2004, DiSantis et al., 2016).

Paper 3: The association between neighborhood type and relative availability of sit-down restaurants and supermarkets in the Twin Cities Region of Minnesota

The third chapter of my dissertation details my examination of the differences of and changes in availability of neighborhood restaurants and food stores over time using as my study area the Twin Cities Region of Minnesota (abbreviated as Twin Cities Region below), an area of nearly three million people living in 186 communities across the seven counties of Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. The neighborhood environment in the Twin Cities Region has become increasingly diverse in physical form and social composition over the past decades (Minneapolis Metropolitan Council, 2015), which provides the opportunity to observe differences of and changes in the distribution of neighborhood food resources over time.

I used a novel concept of the "overall character" to categorize neighborhoods in the Twin Cities Region into six types using 13 neighborhood-level built environmental and sociodemographic factors, which are high density urban core, low-SES inner city, aging suburb, urban, high-income suburb, and suburban edge. As neighborhoods may be systematically different in metropolitan areas by characteristics such as income, race and urbanicity in the U.S., my concept of assessing neighborhoods by defining their "atmosphere" or "image" was inspired by the thought that such qualities are likely important considerations which business owners ponder when deciding whether to open a retail location (Zukin et al., 2017), or the community government keeps in mind when deciding whether to restrict a new business, particularly for small businesses such as fast food restaurants (Carroll and Torfason, 2011). The idea of the "overall character" is also driven by the mixed results generated by previous attempts to categorize neighborhoods using a single construct of neighborhood context (Lytle, 2009), such as income or race. For example, several reports from the literature have indicated that low-income neighborhoods tend to have a greater availability of fast-food restaurants (Powell et al., 2007, Zenk and Powell, 2008), but this has been contradicted by others (Lamichhane et al., 2013, James et al., 2014). In fact, we know little about neighborhoods defined using a more nuanced categorization, which is unfortunate given that the human concept of "neighborhood" is patterned across many interrelated built environment and sociodemographic characteristics (Jones and Huh, 2014). The types of restaurants and food stores available in neighborhoods also vary by built environmental factors such as population density and land use pattern in addition to sociodemographic factors. For example, restaurants and food stores may choose to open outlets even in poor neighborhoods, if residential densities are sufficiently high to produce a demand for their products simply because there is a larger population (Helling and Sawicki, 2003). My way of defining neighborhood type sided with the assertions of those proposing an ecological framework under which the physical environment characteristics which determined where restaurants and food stores are situated interrelate with each other (Story et al., 2008) and do not appear in isolation in neighborhoods (Nelson et al., 2006).

I used a novel dimension of neighborhood food availability to define access to food, which is the relative availability of food outlets, i.e., percentage of sit-down restaurants relative to the total number of sit-down restaurants and fast food restaurants and the percentage of supermarkets relative to the total number of supermarkets, grocery stores, and convenience stores. My work is therefore the first to track both the difference of and the change in the composition of neighborhood food outlets by neighborhood type over a relatively long time period (1993-2011). Given that the neighborhoods represented in my data had increasingly easy access to all foods regardless of neighborhood type over time (Ploeg et al., 2009), perhaps contrary to my expectations I observed a higher relative availability of sit-down restaurants in inner city neighborhoods than I did in either the urban, aging suburb, high-income suburb, or suburban edge neighborhoods in 2011; I did not observe this in 1993 or 2001. My results emphasized the need to examine what factors contributed to the presence of more sit-down restaurants than fast food restaurants in the inner city neighborhoods. My results also suggest the urgent need to examine the implications for those who live in the inner city neighborhoods with an increasing absolute and relative number of sit-down restaurants as sit-down restaurants are a heterogeneous group of restaurants and does not necessarily represent facilities that only sell healthy food options.

REFERENCES

Boone-Heinonen, J., Gordon-Larsen, P., Kiefe, C. I., Shikany, J. M., Lewis, C. E. & Popkin, B. M. 2011a. Fast food restaurants and food stores: Longitudinal associations with diet in young to middle-aged adults: the CARDIA study. *Archives of internal medicine,* 171**,** 1162-1170.

Bucher, T., Collins, C., Rollo, M. E., McCaffrey, T. A., De Vlieger, N., Van der Bend, D., Truby, H. & Perez-Cueto, F. J. 2016. Nudging consumers towards healthier choices: a systematic review of positional influences on food choice. *British Journal of Nutrition,* 115**,** 2252-2263.

Burgoine, T., Forouhi, N. G., Griffin, S. J., Wareham, N. J. & Monsivais, P. 2014. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study. *BMJ,* 348**,** g1464.

Burgoine, T. & Monsivais, P. 2013. Characterising food environment exposure at home, at work, and along commuting journeys using data on adults in the UK. *International Journal of Behavioral Nutrition and Physical Activity,* 10**,** 85.

Carroll, G. R. & Torfason, M. T. 2011. Restaurant Organizational Forms and Community in the US in 2005. *City & Community,* 10**,** 1-24.

Christian, T. & Rashad, I. 2009. Trends in US food prices, 1950–2007. *Economics & Human Biology,* 7**,** 113-120.

Clary, C. M., Ramos, Y., Shareck, M. & Kestens, Y. 2015. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Preventive medicine,* 71**,** 83-87.

Clifton, K. J. 2004. Mobility strategies and food shopping for low-income families: A case study. *Journal of Planning Education and Research,* 23**,** 402-413.

Cummins, S., Flint, E. & Matthews, S. A. 2014. New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity. *Health affairs,* 33**,** 283-291.

Cummins, S., Petticrew, M., Higgins, C., Findlay, A. & Sparks, L. 2005. Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. *Journal of Epidemiology & Community Health,* 59**,** 1035-1040.

DiSantis, K. I., Hillier, A., Holaday, R. & Kumanyika, S. 2016. Why do you shop there? A mixed methods study mapping household food shopping patterns onto weekly routines of black women. *International Journal of Behavioral Nutrition and Physical Activity,* 13**,** 11.

Dubowitz, T., Ghosh-Dastidar, M., Cohen, D. A., Beckman, R., Steiner, E. D., Hunter, G. P., Flórez, K. R., Huang, C., Vaughan, C. A. & Sloan, J. C. 2015. Diet and perceptions change with supermarket introduction in a food desert, but not because of supermarket use. *Health affairs,* 34**,** 1858-1868.

Eating, F. H. 2008. A Planners Guide to Community and Regional Food Planning.

Elbel, B., Moran, A., Dixon, L. B., Kiszko, K., Cantor, J., Abrams, C. & Mijanovich, T. 2015. Assessment of a government-subsidized supermarket in a high-need area on household food availability and children's dietary intakes. *Public health nutrition,* 18**,** 2881-2890.

Ewing, R. & Cervero, R. 2001. Travel and the built environment: a synthesis. *Transportation Research Record,* 1780**,** 87-114.

Forsyth, A., Wall, M., Larson, N., Story, M. & Neumark-Sztainer, D. 2012. Do adolescents who live or go to school near fast-food restaurants eat more frequently from fast-food restaurants? *Health & place,* 18**,** 1261-1269.

Gordon-Larsen, P. 2014. Food availability/convenience and obesity. *Advances in Nutrition: An International Review Journal,* 5**,** 809-817.

Handbury, J., Rahkovsky, I. & Schnell, M. 2015. What drives nutritional disparities? Retail access and food purchases across the socioeconomic spectrum. National Bureau of Economic Research.

Helling, A. & Sawicki, D. S. 2003. Race and residential accessibility to shopping and services. *Housing Policy Debate,* 14**,** 69-101.

James, P., Arcaya, M. C., Parker, D. M., Tucker-Seeley, R. D. & Subramanian, S. 2014. Do minority and poor neighborhoods have higher access to fast-food restaurants in the United States? *Health & place,* 29**,** 10-17.

Jeffery, R. W., Baxter, J., McGuire, M. & Linde, J. 2006a. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition and Physical Activity,* 3**,** 1.

Jones, M. & Huh, J. 2014. Toward a multidimensional understanding of residential neighborhood: A latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. *Health & place,* 27**,** 134-141.

Kerr, J., Frank, L., Sallis, J. F., Saelens, B., Glanz, K. & Chapman, J. 2012. Predictors of trips to food destinations. *International Journal of Behavioral Nutrition and Physical Activity,* 9**,** 58.

Kim, M.-J., Song, S., Park, S. H. & Song, Y. 2015. The association of snack consumption, lifestyle factors, and pediatric obesity with dietary behavior patterns in male adolescents. *Journal of Nutrition and Health,* 48**,** 228-235.

Kyureghian, G., Nayga, R. M. & Bhattacharya, S. 2013. The effect of food store access and income on household purchases of fruits and vegetables: A mixed effects analysis. *Applied Economic Perspectives and Policy,* 35**,** 69-88.

Lamichhane, A. P., Warren, J., Puett, R., Porter, D. E., Bottai, M., Mayer-Davis, E. J. & Liese, A. D. 2013. Spatial patterning of supermarkets and fast food outlets with respect to neighborhood characteristics. *Health & place,* 23**,** 157-164.

Laska, M. N., Graham, D. J., Moe, S. G. & Van Riper, D. 2010a. Young adult eating and food-purchasing patterns: Food store location and residential proximity. *American journal of preventive medicine,* 39**,** 464-467.

Laska, M. N., Hearst, M. O., Forsyth, A., Pasch, K. E. & Lytle, L. 2010b. Neighbourhood food environments: Are they associated with adolescent dietary intake, food purchases and weight status? *Public health nutrition,* 13**,** 1757-1763.

Laxy, M., Malecki, K. C., Givens, M. L., Walsh, M. C. & Nieto, F. J. 2015. The association between neighborhood economic hardship, the retail food environment, fast food intake, and obesity: Findings from the Survey of the Health of Wisconsin. *BMC public health,* 15**,** 237.

Lin, B.-H., Ver Ploeg, M., Kasteridis, P. & Yen, S. T. 2014. The roles of food prices and food access in determining food purchases of low-income households. *Journal of Policy Modeling,* 36**,** 938-952.

Lytle, L. A. 2009. Measuring the food environment: state of the science. *American journal of preventive medicine,* 36**,** S134-S144.

Mercille, G., Richard, L., Gauvin, L., Kestens, Y., Shatenstein, B., Daniel, M. & Payette, H. 2012. Associations between residential food environment and dietary patterns in urban-dwelling older adults: results from the VoisiNuAge study. *Public health nutrition,* 15**,** 2026-2039.

Minneapolis Metropolitan Council 2015. Thrive MSP 2040 Plan: housing policy plan. St. Paul, MN: Metropolitan Counci.

Nelson, M. C., Gordon-Larsen, P., Song, Y. & Popkin, B. M. 2006. Built and social environments: associations with adolescent overweight and activity. *American journal of preventive medicine,* 31**,** 109-117.

Oexle, N., Barnes, T. L., Blake, C. E., Bell, B. A. & Liese, A. D. 2015. Neighborhood fast food availability and fast food consumption. *Appetite,* 92**,** 227-232.

Papas, M. A., Alberg, A. J., Ewing, R., Helzlsouer, K. J., Gary, T. L. & Klassen, A. C. 2007. The built environment and obesity. *Epidemiologic reviews,* 29**,** 129-143.

Popkin, B. M. 2011. Agricultural policies, food and public health. *EMBO reports,* 12**,** 11-18.

Pothukuchi, K. & Kaufman, J. L. 1999. Placing the food system on the urban agenda: The role of municipal institutions in food systems planning. *Agriculture and human values,* 16**,** 213-224.

Powell, L. M., Chaloupka, F. J. & Bao, Y. 2007. The availability of fast-food and full-service restaurants in the United States: associations with neighborhood characteristics. *American journal of preventive medicine,* 33**,** S240-S245.

Richardson, A. S., Boone-Heinonen, J., Popkin, B. M. & Gordon-Larsen, P. 2011. Neighborhood fast food restaurants and fast food consumption: A national study. *BMC Public Health,* 11**,** 1.

Rose, D. & Richards, R. 2004. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public health nutrition,* 7**,** 1081-1088.

Rummo, P. E., Guilkey, D. K., Ng, S. W., Popkin, B. M., Evenson, K. R. & Gordon-Larsen, P. 2017. Beyond Supermarkets: Food Outlet Location Selection in Four US Cities Over Time. *American Journal of Preventive Medicine,* 52**,** 300-310.

Rummo, P. E., Meyer, K. A., Boone-Heinonen, J., Jacobs Jr, D. R., Kiefe, C. I., Lewis, C. E., Steffen, L. M. & Gordon-Larsen, P. 2015. Neighborhood availability of convenience stores and diet quality: Findings from 20 years of follow-up in the Coronary Artery Risk Development in Young Adults study. *American journal of public health,* 105**,** e65-e73.

Sadler, R. C., Gilliland, J. A. & Arku, G. 2013. A food retail-based intervention on food security and consumption. *International journal of environmental research and public health,* 10**,** 3325-3346.

Schwartz, M. B., Just, D. R., Chriqui, J. F. & Ammerman, A. S. 2017. Appetite self-regulation: Environmental and policy influences on eating behaviors. *Obesity,* 25.

Story, M., Kaphingst, K. M., Robinson-O'Brien, R. & Glanz, K. 2008. Creating healthy food and eating environments: policy and environmental approaches. *Annu. Rev. Public Health,* 29**,** 253-272.

Sturm, R. & Hattori, A. 2015. Diet and obesity in Los Angeles County 2007–2012: Is there a measurable effect of the 2008 "Fast-Food Ban"? *Social science & medicine,* 133**,** 205-211.

Thornton, L. E., Pearce, J. R. & Kavanagh, A. M. 2011. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 71.

Ver Ploeg, M. 2010. *Access to affordable and nutritious food: measuring and understanding food deserts and their consequences: report to Congress*, DIANE Publishing.

Volpe, R., Okrent, A. & Leibtag, E. 2013. The effect of supercenter-format stores on the healthfulness of consumers' grocery purchases. *American Journal of Agricultural Economics***,** aas132.

Witten, K., Blakely, T., Bagheri, N., Badland, H., Ivory, V., Pearce, J., Mavoa, S., Hinckson, E. & Schofield, G. 2012. Neighborhood built environment and transport and leisure physical activity: Findings using objective exposure and outcome measures in New Zealand. *Environmental Health Perspectives,* 120**,** 971.

Wrigley, N., Warm, D. & Margetts, B. 2003. Deprivation, diet, and food-retail access: Findings from the Leeds 'food deserts' study. *Environment and Planning A,* 35**,** 151-188.

Zenk, S. N. & Powell, L. M. 2008. US secondary schools and food outlets. *Health & place,* 14**,** 336-346.

Zenk, S. N., Tarlov, E., Wing, C., Matthews, S. A., Jones, K., Tong, H. & Powell, L. M. 2017. Geographic Accessibility Of Food Outlets Not Associated With Body Mass Index Change Among Veterans, 2009–14. *Health Affairs,* 36**,** 1433-1442.

Zukin, S., Lindeman, S. & Hurson, L. 2017. The omnivore's neighborhood? Online restaurant reviews, race, and gentrification. *Journal of Consumer Culture,* 17**,** 459-479.

CHAPTER 1. HOME NEIGHBORHOOD FOOD AVAILABILITY, STREET CONNECTIVITY AND FREQUENCY OF USE OF RESTAURANTS AND FOOD STORES

1.1 Background

The body of literature devoted to how individuals interact with the food environment is growing (Clary et al., 2017), including a focus on the relationship between neighborhood food availability and frequency of use of restaurants and food stores. Studies of neighborhood restaurant/food store availability in relation to frequency of use have produced mixed findings. Several have provided support for a positive association (Boone-Heinonen et al., 2011a, Forsyth et al., 2012, Laxy et al., 2015), whereas others have not (Richardson et al., 2011, Jeffery et al., 2006a, Oexle et al., 2015, Laska et al., 2010b). However, previous studies have focused on the overall frequency of restaurant or food store use without measuring whether the restaurants/food stores that were used were within or outside of study participants' neighborhoods. Neighborhood food availability might be weakly associated with overall frequency of use if most participants used restaurants/food stores in other settings, such as school or workplace neighborhoods (Richardson et al., 2011, Burgoine and Monsivais, 2013). Thus, identifying how neighborhood food outlets are used by neighborhood residents is an important step towards a better understanding of complex dietary behaviors and in developing neighborhood-based environmental strategies to increase healthy diet behaviors.

In addition, few population-based studies have explicitly addressed whether street connectivity plays a role in the association between neighborhood restaurant/food store availability and use of restaurants/food stores (Shannon, 2016). Previous studies have indicated that people in highly connected (or walkable) neighborhoods were more likely to participate in transport walking (e.g., walking to work or a grocery store) than those in poorly connected neighborhoods (Witten et al., 2012, Ewing and Cervero, 2001), as connectivity affects the directness of travel and the number of alternative routes (Thornton et al., 2011). Street connectivity, reflecting the convenience of accessing food resources, might modify the strength of the relationship between neighborhood restaurant/food store availability and frequency of using those food resources.

I therefore used data on the self-reported frequency of using fast food restaurants/sit-down restaurants/grocery stores in participants' home neighborhoods from the Coronary Artery Risk Development in Young Adults (CARDIA) study for exam year 2005-2006. Using geographically matched restaurant and food store locations and street information, I estimated the associations between the GIS-measured number of neighborhood fast food restaurants/sit-down restaurants/grocery stores and self-reported frequency of using fast food restaurants/sit-down restaurants/grocery stores in respondents' neighborhoods. To address the potential role of neighborhood fast food restaurants/sit-down restaurants/grocery stores in disparities in street connectivity, I also examined how such associations differed by measured neighborhood street connectivity.

1.2 Methods

CARDIA is a prospective cohort study examining the development of cardiometabolic disease in 5115 white or black U.S. adults aged 18 to 30 years who were recruited to attain an approximately balanced representation of age (18-24 years or 25-30 years), race (White or Black), gender, and education (≤ high school versus > high school) from 4 metropolitan study centers (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA) in 1985-1986 (Friedman et al., 1988). I used data on 3549 participants in exam year 2005-2006, in which CARDIA surveyed people how they used neighborhood food outlets through a neighborhood environment questionnaire. I included only the participants who completed the neighborhood environment questionnaire (n=3539). I excluded 576 participants who did not report the question about the existence of a food outlet in their neighborhood. I also excluded 42 participants due to invalid skip patterns. In addition, I excluded 61 participants due to missing covariate information, resulting in a final sample of n=2860 (Figure 1.1). I used a geographic information system to link these three types of food outlets, street, and US Census data to CARDIA participants' residential addresses.

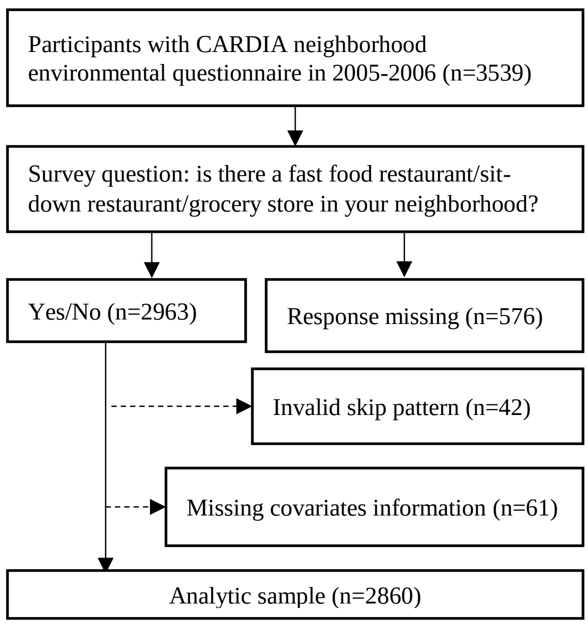

Figure 1.1. Selection of participants for a study of frequencies of using food outlets, CARDIA, 2005-2006

1.2.1 Outcome: self-reported frequency of using neighborhood fast food restaurants/sit-down restaurants/grocery stores

I used the CARDIA neighborhood environment questionnaire, which elicited information on the frequency of use of each type of neighborhood food outlet (fast food restaurants, sit-down restaurants, or grocery stores) separately in each participant's neighborhood. If the participant answered "yes" to the first question ("Is there a fast food restaurant/sit-down restaurant/grocery store in your neighborhood (q1)?”), they were then asked a second question (“In the past year, did you use a fast food restaurant/sit-down restaurant/grocery store in your neighborhood (q2)?”; if the answer to this question was “yes,” then the participant was asked a third question (“How often did you use a fast food restaurant/sit-down restaurant/grocery store (q3)?”), with response options of “more than once per week,” “weekly,” “monthly,” and “yearly.” The final analytic samples comprised 2007, 2122, and 2191 participants who explicitly reported the frequency of using neighborhood fast food restaurants, sit-down restaurants, and grocery stores, respectively.

1.2.2 Exposures: GIS-measured neighborhood fast food restaurants, sit-down restaurants, and grocery stores and neighborhood street connectivity

The CARDIA neighborhood environment questionnaire defined neighborhood as an area within a 10- to 15-minute walk from the participant’s home. I therefore used the 1-km Euclidean buffer around each participant’s geocoded residential location to operationalize neighborhood because the distance that can be travelled from home by walk in 10-15 minutes is 0.7-1.1 kilometers (given a walking speed of approximately 4.4 kilometers per hour) (Rodríguez and Joo, 2004). I defined number of neighborhood fast food restaurants, sit-down restaurants, and grocery stores within the 1-km buffer. I calculated the number of fast food restaurants/sit-down restaurants/grocery stores by geocoding the food outlet records retrieved from the D&B, a commercial dataset of U.S. business records. I defined food outlets according to their primary 8-digit SIC codes. I used the link-to-node ratio (number of links divided by the number of nodes) within the 1-km buffer to describe neighborhood street connectivity (Berrigan et al., 2010). A higher ratio indicates higher connectivity. I obtained the road network maps (interstate highways and access ramps excluded) from the ESRI Data and Maps StreetMap North America for 2010. Additional information on my food outlet classification and the detailed method I used to produce exposures can be found in Table A1-1 in Appendix 1-1 and Appendix 1-2.

1.2.3 Covariates

I used self-reported individual-level sociodemographic information collected at baseline exam year (1985-1986) using the CARDIA standardized questionnaire, including race (Black, White), gender, and age. I used self-reported individual-level sociodemographic and other information collected at exam year (2005-2006) using the CARDIA sociodemographic and neighborhood environment questionnaires, including current educational attainment (≤ high school, > high school), family income, household size, employment status (employed, not employed), marital status (married, not married), and reasons for moving to or staying in the current neighborhood. Additionally, I used five GIS-measured neighborhood-level variables to account for the contextual influences of other neighborhood-level built environment and sociodemographic characteristics, including population density within the 1-km buffer, SES deprivation factor score (defined as the first factor score from a principal components analysis of four census indicators of socioeconomic status) in a participant's home census tract, density of vacant housing units in a participant's home census block group, total number of neighborhood fast food restaurants and sit-down restaurants within the 1-km buffer, total number of neighborhood supermarkets and convenience stores within the 1-km buffer. The detailed methods I used to produce these covariates can be found in Appendix 1-2.

1.2.4 Statistical analysis

I used a separate set of models for each type of outlet (fast food restaurants, sit-down restaurants, grocery stores) to predict the self-reported frequency of use of each. Not all participants perceived all three types of neighborhood food outlets and therefore they did not report frequency of use for outlets not perceived. However, it is likely that participant reporting not having an outlet may be different than those having an outlet, leading to bias in the estimated coefficients of frequency of use. To address this, I used propensity score: first I estimated the probability of whether or not a participant reported having a given food outlet in their neighborhood and second I estimated frequency of use of given neighborhood outlet by including the estimated step 1 probability as one of covariates (Cuddeback et al., 2004, Rosenbaum and Rubin, 1983).

For each food outlet type, the first equation is a random intercept mixed effects logistic regression that estimates the probability of perceiving at least one neighborhood food outlet across the full sample (n=2860). The second equation is a random intercept mixed effects generalized ordered logistic regression, also known as a mixed effects proportional odds model, that estimates the participant's self-reported frequency of using the food outlet for participants who reported perceiving at least one outcome food outlet (fast food restaurant: n=2007; sit-down restaurant: n=2122; grocery store: n=2191). I adjusted the coefficients from the second equation by adding the predicted probability of perceiving at least one outcome food outlet from the first equation into the model as a covariate in the second equation.

I further adjusted for the following self-reported individual-level covariates in all models, as suggested by previous studies: family income (Boone-Heinonen et al., 2011a, Forsyth et al., 2012, Moore et al., 2009, Richardson et al., 2011), race (Boone-Heinonen et al., 2011a, Forsyth et al., 2012, Richardson et al., 2011, Moore et al., 2009), gender (Forsyth et al., 2012, Laska et al., 2010b, Roda et al., 2016), age (Boone-Heinonen et al., 2011a, Roda et al., 2016), employment status (Barnes et al., 2015, Richardson et al., 2011), and whether neighborhood food environment (grocery stores, restaurants, corner stores) was one of the most important reasons for moving to/staying in the neighborhood. I also adjusted for the following neighborhood-level covariates in all models: population density within the 1-km buffer (Boone-Heinonen et al., 2011a, Roda et al., 2016, Barnes et al., 2015), SES deprivation factor score in a participant's home census tract (Roda et al., 2016), and density of vacant housing units in a participant's home census block group. In addition to the predicted probability of perceiving at least one food outlet, I adjusted for three self-reported individual-level covariates (educational attainment (Boone-Heinonen et al., 2011a), household size (Boone-Heinonen et al., 2011a), and marital status (Boone-Heinonen et al., 2011a)) in the frequency of use equations since I assumed that these covariates were associated with the self-reported frequency of use but not necessarily associated with perceiving at least one neighborhood food outlet. To adjust for presence of other complementary neighborhood food outlets, I controlled for other restaurants and other food shopping sources in each of the models (e.g., I controlled for the number of GIS-measured sit-down restaurants in the fast food restaurant model).

For the possible modification effect of neighborhood street connectivity, I included an interaction term between the GIS-measured number of each type of food outlet and the link-to-node ratio measure. I ran the collinearity diagnostics and mean-centered exposures to address potential collinearity concerns. I used the *collin* command in STATA 14 (StataCorp, College Station, TX) to run the collinearity diagnostics. Additional information about collinearity diagnostics can be found in Appendix 1-4. As none of the interaction term was significant in the fast food restaurant, sit-down restaurant or the grocery store models, I did not include the interaction terms in the final frequency of use models.

I tested the proportional odds assumption for the random intercept mixed effects ordinal frequency of use of neighborhood fast food restaurants/sit-down restaurants/grocery stores models to ensure the relationship between all pairs of outcome groups is the same and that there is only one set of coefficients. The results of the likelihood ratio chi-squared test suggested some variables in the ordinal frequency of use models violated the assumption, which are shown as bolded text in Table A1-4e and A1-4f in Appendix 1-4. I relaxed the proportional odds assumptions for the variables that violated the assumption. For example, in the 1-km buffer models, education attainment and household size in the ordinal frequency of use of neighborhood fast food restaurants model violated the assumption; I therefore relaxed the assumption for education attainment and household size, which obtained four coefficients which indicated the effects of switching from *no use* to *use yearly*, from *use yearly* to *use monthly*, from *use monthly* to *use weekly*, and from *use weekly* to *use more than once per week*, respectively. I used Stata 14.0 for regression analyses (*xtmelogit* for mixed effects logistic regression, *gllamm*, *threshold*, and *test* for testing the proportional odds assumption and mixed effects generalized ordered logistic regression).

1.2.5 Sensitivity testing

Since participants may misestimate the geographic boundaries achievable by a 10 to 15 minute walk from home(Moore et al., 2008), I tested whether my results were sensitive with respect to distances by using also a 3-km buffer. In addition, I examined the similarity of individual-level and neighborhood-level characteristics between individuals who perceived at least one outcome food outlet versus those who did not to examine this source of selection bias using kernel density plots of each to compare similarity between these two groups. The overlap of the plotted curves between these two groups reflected the balance of neighborhood-level and individual-level characteristics between these two groups. Additional information about the balance tests can be found in Appendix 1-4.

1.3 Results

More than two-thirds of the participants reported that they perceived at least one neighborhood fast food restaurant/sit-down restaurant/grocery store. More than half of the participants who perceived that they had these outlets in their neighborhoods reported that they used the neighborhood fast food restaurants/sit-down restaurants while around three-fourths reported that they used the neighborhood grocery stores (Table 1.1).

Table 1.1. Characteristics of CARDIA study sample

Characteristics	N=2860
Perceived presence of at least one outcome food outlet	
Is there a fast food restaurant in your neighborhood, %	
Yes	69.8
Is there a sit-down restaurant in your neighborhood, %	
Yes	74.1
Is there a grocery store in your neighborhood, %	
Yes	76.5
Self-reported frequency of use of neighborhood food outlets	
Fast food restaurant, %	
Never use	13.2
Once per year	5.9
Monthly	28.5
Weekly	15.1
More than once per week	7.5
No response [a]	29.8
Sit-down restaurant, %	
Never use	15.0
Once per year	10.7
Monthly	32.6
Weekly	11.5
More than once per week	4.3
No response [a]	25.8
Grocery store, %	
Never use	1.9
Once per year	1.7
Monthly	12.3
Weekly	35.3
More than once per week	25.4
No response [a]	23.4
GIS-measured neighborhood food availability measures	
Number of food outlets within the 1-km buffer, count, mean (SD)	
Fast food restaurant [b]	6.2 (13.1)
Sit-down restaurant [b]	5.6 (15.9)
Grocery store [b]	3.9 (8.3)
Supermarkets and convenience stores [b]	4.0 (5.2)
GIS-measured neighborhood street connectivity measure	
Link-to-node ratio within the 1-km buffer, mean (SD)	1.6 (0.2)
Other GIS-measured neighborhood environmental measures	
Population density within the 1-km buffer: 1000 person/km^2, mean (SD)	2.4 (3.1)
Neighborhood SES deprivation based on participants' home census tract, mean (SD)	0.4 (1.1)
Density of vacant housing units in participants' home census block group: 100 housing units/km^2, mean (SD)	1.3 (3.8)

Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood	
Education> high school, %	62.4
Family income, 1000 $, mean (SD)	74.3 (41.0)
Black, %	44.3
Female, %	56.5
Household size, person, mean (SD)	3.0 (1.5)
Age, years, mean (SD)	45.2 (3.6)
Employed, %	82.4
Married, %	56.8
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood, %	25.2
Study center	
Birmingham	24.1
Chicago	23.3
Minneapolis	24.4
Oakland	28.3

Abbreviations: SD: standard deviation; km: kilometer.
Notes [a] Participants did not answer the frequency of use because they reported that they did not perceive a neighborhood fast food restaurant/sit-down restaurant/grocery store and therefore they are not required to answer the following question—frequency of using neighborhood fast food restaurants/sit-down restaurants/grocery stores.
[b] See Table A1-1 in Appendix 1-1 for food outlet classification using SIC codes.

1.3.1 Associations between GIS-measured neighborhood food availability, neighborhood street connectivity, and perceiving at least one neighborhood food outlet type

After adjusting for covariates, I observed that: 1) the probability of perceiving at least one neighborhood fast food restaurant was positively associated with the GIS-measured number of fast food restaurants within the 1-km buffer and inversely associated with the link-to-node ratio within the same buffer; 2) the probability of perceiving at least one neighborhood sit-down restaurant was positively associated with the GIS-measured number of sit-down restaurants within the same buffer (Table 1.2-1.4).

1.3.2 Associations between GIS-measured neighborhood food availability, neighborhood street connectivity, and self-reported frequency of use of neighborhood fast food restaurants/sit-down restaurants/grocery stores

After adjusting for covariates, I observed that: 1) the self-reported frequency of using neighborhood fast food restaurants was inversely associated with the link-to-node ratio within the 1-km buffer; 2) the self-reported frequency of using neighborhood sit-down restaurants was positively associated with the GIS-measured number of sit-down restaurants within the same buffer; 3) the self-reported frequency of using neighborhood grocery stores (i.e., switching from never use to use yearly) was inversely associated with the link-to-node ratio within the 1-m buffer (Table 1.2-1.4).

Table 1.2. Associations between GIS-measured neighborhood fast food restaurant availability, neighborhood street connectivity and self-reported frequency of use of neighborhood fast food restaurants

GIS-measured exposure	First-step model: perceiving at least one neighborhood fast food restaurant [a] (full sample)	Second-step model: self-reported frequency of use of neighborhood fast food restaurants [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2007)
Number of fast food restaurants	**1.05 (1.03, 1.07)**	1.01 (0.99, 1.03)
Link-to-node ratio	**0.61 (0.38, 0.97)**	**0.42 (0.26, 0.68)**

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates significant association (P <.05).
Notes: [a] Estimated coefficients of perceiving at least one fast food restaurant in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.
[b] Estimated coefficients of frequency of use of neighborhood fast food restaurants within the participants who perceived at least one fast food restaurant in the participant's neighborhood were adjusted for number of neighborhood sit-down restaurants, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one fast food restaurant in the participant's neighborhood.

Table 1.3. Associations between GIS-measured neighborhood sit-down restaurant availability, neighborhood street connectivity and self-reported frequency of use of neighborhood sit-down restaurants

GIS-measured exposure	First-step model: perceiving at least one neighborhood sit-down restaurant [a] (full sample)	Second-step model: self-reported frequency of use of neighborhood sit-down restaurants [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2122)
Number of sit-down restaurants	**1.05 (1.02, 1.07)**	**1.02 (1.00, 1.04)**
Link-to-node ratio	1.53 (0.93, 2.53)	0.84 (0.52, 1.35)

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates significant association (P <.05).

Notes: [a] Estimated coefficients of perceiving at least one sit-down restaurant in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.

[b] Estimated coefficients of frequency of use of neighborhood sit-down restaurants within the participants who perceived at least one sit-down restaurant in the participant's neighborhood were adjusted for number of neighborhood fast food restaurants, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one sit-down restaurant in the participant's neighborhood.

Table 1.4. Associations between GIS-measured neighborhood grocery store availability, neighborhood street connectivity and self-reported frequency of use of neighborhood grocery stores

GIS-measured exposure	First-step model: perceiving at least one neighborhood grocery store (full sample) [a]	Second-step model: self-reported frequency of use of neighborhood grocery stores [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2191)
Number of grocery stores	0.98 (0.95, 1.01)	-0.13 (-0.32, 0.06) [c] 0.03 (-0.02, 0.09) [c] 0.01 (-0.01, 0.04) [c] 0.02 (-0.00, 0.04) [c]
Link-to-node ratio	0.81 (0.49, 1.35)	**-2.26 (-4.52, -0.01)** [c] 0.59 (-0.54, 1.72) [c] 0.14 (-0.44, 0.72) [c] -0.03 (-0.52, 0.47) [c]

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates statistically significant at $p < 0.05$.

Notes: [a] Estimated coefficients of perceiving at least one grocery store in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.

[b] Estimated coefficients of frequency of use of neighborhood grocery stores within the participants who perceived at least a grocery store in the participant's neighborhood only were adjusted for number of neighborhood supermarkets and convenience stores, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one grocery store in the participant's neighborhood.

[c] The number of grocery stores and link-to-node ratio within the 1-km buffer had four coefficients because the variables violated the proportional odds assumption; the four coefficients indicated the effects of the exposures switching from no use to use yearly, from use yearly to use monthly, from use monthly to use weekly, and from use weekly to use more than once per week, respectively

1.3.3 Sensitivity testing

After adjusting for covariates, results for the self-reported frequency of use model using the 3-km buffer were largely consistent with those using the 1-km buffer but the results for perceived presence of food outlets models showed some inconsistencies. The probability of perceiving at least one neighborhood fast food restaurant was positively associated with the GIS-measured number of fast food restaurants within the 1-km buffer and the link-to-node ratio within the 1-km buffer but not associated with those within the 3-km buffer. The probability of perceiving at least one sit-down restaurant was associated with the GIS-measured number of sit-down restaurants within the 1-km buffer but not within the 3-km buffer. Based on these results, my use of the 1-km buffer for my measures of GIS-measured exposure may have been more aligned with the participants' perception of their neighborhoods than was the use of those measures within the 3-km buffer, especially fast food restaurants. Regression results using the 3-km buffer can be found in Table A1-3a, A1-3b and A1-3c in Appendix 1-3.

I also found that the individuals who reported perceiving at least one fast food restaurant/sit-down restaurant/grocery store had similar individual-level and neighborhood-level characteristics than those who did not. The plotted curves (showing overlap between these two groups) are shown in Figure A1-4 in Appendix 1-4. This ameliorated concerns regarding off-support inference (Oakes and Johnson, 2006) and allowed me to include all the participants who explicitly reported the frequency of use into the frequency of use models based on the evidence that they did not differ greatly from those who reported that they did not perceive at least one outcome food outlet (and therefore did not report the frequency of use).

1.4 Discussion

I found some evidence of cross-sectional associations between GIS-measured neighborhood food availability, neighborhood street connectivity, and the self-reported frequency of using neighborhood food outlets in a large, diverse cohort of middle-aged adults. Although some of my results highlight the association between food outlets and street connectivity in the immediate home neighborhood in relation to use of neighborhood food outlets, they also underscore the complexity underlying the relationship between the availability of different types of neighborhood food outlets and use of neighborhood food outlets.

GIS-measured availability of neighborhood fast food restaurants were not associated with the self-reported frequency of use of such neighborhood outlets. My findings agree with the previous literature suggesting that availability of neighborhood fast food restaurants was not related to the overall frequency of use of fast food restaurants (Richardson et al., 2011, Jeffery et al., 2006b, Oexle et al., 2015, Laska et al., 2010b). One previous study identified a significant higher overall frequency of use of fast food restaurants among adolescent males living in neighborhoods with high numbers of such restaurants (Forsyth et al., 2012). It is possible that there is a relationship between spatial distribution (e.g., clustering near home, schools or highway off-ramps) of neighborhood fast food restaurants (beyond the simple count of such restaurants) and use of such outlets, particularly in high-density areas shown to have clustering of fast food restaurants (Austin et al., 2005), which should be examined in the future.

Participants in neighborhoods with an (GIS-measured) greater number of sit-down restaurants tended to report that they used such restaurants more frequently. A previous study similarly indicated such an association but used the overall frequency of eating at all restaurants other than fast food restaurants (defined as those selling quick service burger, roast beef, and pizza parlor) as the outcome (Jeffery et al., 2006a). My results suggest a greater concern with respect to the use of sit-down restaurants because such restaurants are a heterogeneous group. Although only a small proportion of my study sample used neighborhood sit-down restaurants on a weekly or more frequent basis, this proportion could be greater in neighborhoods with a greater number of sit-down restaurants, which should be examined in the future.

Neighborhood street connectivity measured the directness of possible within-neighborhood routes and the number of optional routes in the home neighborhood from which study participants could choose. The results for street connectivity, showing a negative association with fast food restaurant/grocery store use, are intriguing. Future research should examine and attempt to disentangle the causal pathways for this result.

Consistent with a previous study (Roda et al., 2016), a higher number of neighborhood fast food restaurants and of sit-down restaurants increased participants' awareness of such neighborhood food outlets. However, I did not find such an association for grocery stores. This finding suggests that increasing the number of grocery stores alone is unlikely to change resident's awareness of grocery store availability. It might be that related interventions that assist residents in recognizing, mapping, and sharing information about the types of stores in residential neighborhoods can help with awareness (Horowitz et al., 2004). It is also possible that there is measurement error in the definition of food outlets, for example if participants had different ideas about how to define a grocery store than researchers (Caspi et al., 2012).

My study had several limitations. My assumption that people use food outlets partly based on their awareness of such neighborhood outlets may be inappropriate. Thus, perception ("I'm aware of the food outlet") may not translate to intention ("since I'm aware of the presence of the food outlet, and other conditions (e.g., time, money) are also met, I will eat there"). Second, my measures of frequency of using food outlets were derived from self-reports, which are prone to recall bias and other reporting errors. In the case of grocery stores, people might forget to report small purchases (for example, beverages or a bag of chips) and therefore underestimate frequency of their use. Third, frequency of use failed to indicate the type and amount of food purchased in restaurants and food stores, which is difficult to link to diet-related outcomes; this issue is probably severe with respect to understanding the use of grocery stores because I cannot ascertain what type of food (sweetened beverage or fresh fruits) was purchased in the grocery stores (although assuming I knew the type of the food purchased in fast food restaurants was equally bold). Fourth, people might incorrectly include the use of food outlets outside of the 10-15 minute walk area. I addressed this concern in my sensitivity analysis, using 3-km buffers, finding similar results as with the 1-km buffer. Fifth, my analysis may have omitted important factors that explain residential selection and use of food outlets. Sixth, I noted that CARDIA participants were recruited from urban environments, which may limit the generalizability of my findings to rural samples (Richardson et al., 2011). Also, the Dun & Bradstreet food business record data may have contained location errors, but these errors were probably random in nature and small.

1.5 Conclusions

My findings suggest that individuals who have more sit-down restaurants in their neighborhoods report more frequent use of such restaurants than those whose neighborhood contains fewer such restaurants. My findings also suggest that individuals who live in neighborhoods with greater neighborhood street connectivity report less use of neighborhood fast food restaurants than those who live in neighborhoods with less neighborhood street connectivity. The results for street connectivity, showing an inverse association with fast food restaurant and grocery store use, are intriguing and point to future research to disentangle the causal pathways for this result.

REFERENCES

Austin, S. B., Melly, S. J., Sanchez, B. N., Patel, A., Buka, S. & Gortmaker, S. L. 2005. Clustering of fast-food restaurants around schools: a novel application of spatial statistics to the study of food environments. *American Journal of Public Health,* 95**,** 1575-1581.

Barnes, T. L., Bell, B. A., Freedman, D. A., Colabianchi, N. & Liese, A. D. 2015. Do people really know what food retailers exist in their neighborhood? Examining GIS-based and perceived presence of retail food outlets in an eight-county region of South Carolina. *Spatial and spatio-temporal epidemiology,* 13**,** 31-40.

Berrigan, D., Pickle, L. W. & Dill, J. 2010. Associations between street connectivity and active transportation. *International journal of health geographics,* 9**,** 20.

Boone-Heinonen, J., Gordon-Larsen, P., Kiefe, C. I., Shikany, J. M., Lewis, C. E. & Popkin, B. M. 2011a. Fast food restaurants and food stores: Longitudinal associations with diet in young to middle-aged adults: the CARDIA study. *Archives of internal medicine,* 171**,** 1162-1170.

Burgoine, T. & Monsivais, P. 2013. Characterising food environment exposure at home, at work, and along commuting journeys using data on adults in the UK. *International Journal of Behavioral Nutrition and Physical Activity,* 10**,** 85.

Caspi, C. E., Kawachi, I., Subramanian, S., Adamkiewicz, G. & Sorensen, G. 2012. The relationship between diet and perceived and objective access to supermarkets among low-income housing residents. *Social science & medicine,* 75**,** 1254-1262.

Clary, C., Matthews, S. A. & Kestens, Y. 2017. Between exposure, access and use: Reconsidering foodscape influences on dietary behaviours. *Health & place,* 44**,** 1-7.

Cuddeback, G., Wilson, E., Orme, J. G. & Combs-Orme, T. 2004. Detecting and statistically correcting sample selection bias. *Journal of Social Service Research,* 30**,** 19-33.

Ewing, R. & Cervero, R. 2001. Travel and the built environment: a synthesis. *Transportation Research Record,* 1780**,** 87-114.

Forsyth, A., Wall, M., Larson, N., Story, M. & Neumark-Sztainer, D. 2012. Do adolescents who live or go to school near fast-food restaurants eat more frequently from fast-food restaurants? *Health & place,* 18**,** 1261-1269.

Friedman, G. D., Cutter, G. R., Donahue, R. P., Hughes, G. H., Hulley, S. B., Jacobs, D. R., Liu, K. & Savage, P. J. 1988. CARDIA: Study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology,* 41**,** 1105-1116.

Horowitz, C. R., Colson, K. A., Hebert, P. L. & Lancaster, K. 2004. Barriers to buying healthy foods for people with diabetes: evidence of environmental disparities. *American Journal of Public Health,* 94**,** 1549-1554.

Jeffery, R. W., Baxter, J., McGuire, M. & Linde, J. 2006a. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition and Physical Activity,* 3**,** 1.

Jeffery, R. W., Baxter, J., McGuire, M. & Linde, J. 2006b. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition and Physical Activity,* 3**,** 2.

Laska, M. N., Hearst, M. O., Forsyth, A., Pasch, K. E. & Lytle, L. 2010b. Neighbourhood food environments: Are they associated with adolescent dietary intake, food purchases and weight status? *Public health nutrition,* 13**,** 1757-1763.

Laxy, M., Malecki, K. C., Givens, M. L., Walsh, M. C. & Nieto, F. J. 2015. The association between neighborhood economic hardship, the retail food environment, fast food intake, and obesity: Findings from the Survey of the Health of Wisconsin. *BMC public health,* 15**,** 237.

Moore, L. V., Roux, A. V. D. & Brines, S. 2008. Comparing perception-based and geographic information system (GIS)-based characterizations of the local food environment. *Journal of Urban Health,* 85**,** 206-216.

Moore, L. V., Roux, A. V. D., Nettleton, J. A., Jacobs, D. R. & Franco, M. 2009. Fast-food consumption, diet quality, and neighborhood exposure to fast food the multi-ethnic study of atherosclerosis. *American journal of epidemiology,* 170**,** 29-36.

Oakes, M. & Johnson, P. 2006. Propensity score matching for social epidemiology. Chapter 15 in Methods in Epidemiology, eds. Oakes, MJ and Kaufman, JS New York. John Wiley & Sons, Inc.

Oexle, N., Barnes, T. L., Blake, C. E., Bell, B. A. & Liese, A. D. 2015. Neighborhood fast food availability and fast food consumption. *Appetite,* 92**,** 227-232.

Richardson, A. S., Boone-Heinonen, J., Popkin, B. M. & Gordon-Larsen, P. 2011. Neighborhood fast food restaurants and fast food consumption: A national study. *BMC Public Health,* 11**,** 1.

Roda, C., Charreire, H., Feuillet, T., Mackenbach, J., Compernolle, S., Glonti, K., Ben Rebah, M., Bárdos, H., Rutter, H. & McKee, M. 2016. Mismatch between perceived and objectively measured environmental obesogenic features in European neighbourhoods. *Obesity reviews,* 17**,** 31-41.

Rodríguez, D. A. & Joo, J. 2004. The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment,* 9**,** 151-173.

Rosenbaum, P. R. & Rubin, D. B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70**,** 41-55.

Shannon, J. 2016. Beyond the supermarket solution: Linking food deserts, neighborhood context, and everyday mobility. *Annals of the American Association of Geographers,* 106**,** 186-202.

Thornton, L. E., Pearce, J. R. & Kavanagh, A. M. 2011. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 71.

Witten, K., Blakely, T., Bagheri, N., Badland, H., Ivory, V., Pearce, J., Mavoa, S., Hinckson, E. & Schofield, G. 2012. Neighborhood built environment and transport and leisure physical activity: Findings using objective exposure and outcome measures in New Zealand. *Environmental Health Perspectives,* 120**,** 971.

CHAPTER 2. BUILT ENVIRONMENT AND THE PURCHASE OF FRUITS AND VEGETABLES IN UNITED STATES HOUSEHOLDS

2.1 Introduction

Given that the causal links between built environment and diet-related behaviors are not well-established (Mayne et al., 2015), food purchasing behaviors (e.g., what food is purchased, the amount of food purchased) are an increasingly-frequent topic of epidemiologic and intervention studies, since they may represent mediating factors linking the presence of food outlets in the built environment and behaviors of food consumption (Cobb et al., 2015, Cummins et al., 2014, Appelhans et al., 2017, Minaker et al., 2013). Studies on questions of whether people living in neighborhoods with a greater availability of food stores (e.g., supermarkets, supercenter) tend to purchase more healthy food (e.g., quantity, expenditure, healthfulness score) produced mixed results. The findings reported in several have provided support for a significant association (Weatherspoon et al., 2013, Volpe et al., 2013, Handbury et al., 2015), but those of others did not (Kyureghian et al., 2012, Lin et al., 2014). For example, Handbury et al. (2015) observed that the concentration of retail food stores was positively associated with the healthfulness of household food purchases in the United States (U.S.), yet Kyureghian et al. (2012) observed that the densities of supermarkets in metropolitan areas in the U.S. did not have significant effect on household fruit and vegetable purchases. These mixed results have disturbing implications for spatial policies and programs that are rapidly being adopted in the U.S. that are aimed at improving food access and healthy behaviors (Sturm and Hattori, 2015, Elbel et al., 2015, Cummins et al., 2014).

Previous studies focused on one key component of built environment related to food purchase—the availability of neighborhood supermarkets. Rarely the interconnected and broad built environment is considered (Kerr et al., 2012, Rose and Richards, 2004, Laska et al., 2010a). One theoretical consideration underpinning such work is that households may have greater access to nutritious food items (e.g., fruit, vegetable, whole grain) in a region if they have greater opportunities to use food stores outside of their immediate neighborhood (Widener et al., 2013, Kerr et al., 2012). Conversely, residents living in neighborhoods with greater availability of convenience stores may purchase greater amount of less nutritious food because the availability of energy-dense foods is high relative to that of healthier alternatives (Rose et al., 2009). Thus, identifying how the interconnected and broad built environment (e.g., regional destination accessibility, availability of neighborhood convenience stores) is related to food purchasing behaviors is an important step towards a better understanding of how people use food stores and in developing environmental-based strategies to increase purchasing nutritious food items. Using a national-level dataset may permit the examination of food purchasing behaviors across regions or urban areas, thus increasing generalizability and enabling the extrapolation of the findings to other areas (Laska et al., 2010a).

Using the information on self-reported expenditures for fresh fruits and vegetables by household available from the Nielsen Homescan Consumer Panel Dataset for 2010 (Nielsen, 2016), I have provided the first national study of the cross-sectional associations between built environment and food purchases in a population of 22,448 households in 378 metropolitan statistical areas (abbreviated as MSA below) in the continental U.S. I focused on a range of built environmental factors that facilitate or hinder food purchasing behaviors, such as availability of neighborhood convenience stores, regional destination accessibility, availability of neighborhood destinations, neighborhood destination diversity, and neighborhood street connectivity, in addition to the commonly examined availability of neighborhood supermarkets. Through treating built environment as a multifaceted physical environment, I contribute to an enhanced understanding of the potential connection between modifications that urban planners and policy makers can make to the built environment and to healthy food purchasing behavior.

2.2 Methods

2.2.1 Sample

Nielsen's National Homescan Consumer Panel Dataset (abbreviated as *Nielsen data* below) is an ongoing nationally representative survey of between 40,000 and 60,000 U.S. households that captures household purchases of food and beverage items (Nielsen, 2016). I derived the Nielsen food purchase data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Nielsen households (abbreviated as *households* below) reported food purchases in retailing stores such as warehouse club stores (Costco and Sam's Club); mass merchandisers and supercenters (Walmart and Target); grocery stores; convenience, drug, and dollar stores; ethnic and specialty stores (Compare Foods and Whole Foods Market); and others (department stores and book stores) (Stern et al., 2015). Participating households were given barcode scanners and household members scanned the barcodes on all purchased foods and beverages after every shopping trip. Households also recorded the quantity of items purchased, the price of the item, the location where the item was purchased, and when the item was purchased, etc. I included the n=27,422 magnet households in the total of n=60,658 households available for 2010. Magnet households reported non-standard Universal Product Code (UPC) products (Nielsen, 2016), which included random weighted (loose) items such as fruits, vegetables, meats, and in-store baked goods (Einav et al., 2008, Oster, 2015, Allcott et al., 2017, Zhen et al., 2009), in addition to the standard-UPC products they reported (Oster, 2015, Allcott et al., 2017). Focusing on magnet households provided more information on how people used grocery retailers to purchase fresh fruits and vegetables. I excluded 4,559 households outside the MSAs, 131 households due to missing covariate information, and 284 households with extremely low or high values for expenditures on both fruits and vegetables (below the 2^{nd} percentile or above the 98^{th} percentile of the expenditure for fruits or vegetables), resulting in a final sample of n=22,448 households. See Appendix 2-1 for the sample construction. The number of households in each MSA ranged from 2 (Farmington, NM) to 1,095 (New York- Newark-Jersey City, NY-NJ-PA) with a median value of 151. I then linked the self-reported household-level food purchase and sociodemographic information from the Nielsen data to the built environment characteristics in 2010 to study the cross-sectional association between built environment and food purchase.

2.2.2 Outcome variables

The type of self-reported food purchased I considered included fresh fruits (abbreviated as *fruits* below) and fresh vegetables (abbreviated as *vegetables* below). Fresh foods were defined as those that have not undergone any processing and are therefore in their raw state. To properly classify self-reported non-magnet expenditures, I used the departmental category (fresh produce) and the product module description (e.g., fresh apple, fresh lettuce) to identify fruits or vegetables; for the self-reported magnet expenditures, I used the departmental category (magnet) and brand description (e.g., reference card fruits, reference card vegetables) to identify fruits or vegetables respectively. I then calculated the self-reported expenditures on fruits and vegetables (separately) as the sum of standard-UPC products and non-standard-UPC products by magnet household for 2010, then used these estimates to partially address the potential issue of random purchasing behaviors (i.e., impulsive purchases) in a short observational period (e.g., weekly or monthly) (DiSantis et al., 2016). Although some Nielsen households (e.g., Asian and Hispanic households or households with employed female heads of household) tended to underreport the purchase of fresh fruits and vegetables (Zhen et al., 2009) because many fresh fruits and vegetables were not barcoded and therefore cannot be scanned, I did not find any evidence in previous literature that such underestimation is systematically biased by built environmental factors. One study concluded that the degree of measure error found in Homescan was comparable to that found in other commonly used economic datasets (Einav et al., 2008). I used expenditure values rather than weight values because magnet household only reported expenditures. See Appendix 2-2 for the details of developing outcome variables.

2.2.3 Exposure variables

As Nielsen only disclosed the zip code tabulation area in which the household resided, I used the centroid of the zip code tabulation area as a proxy for the exact residential location of the household. I characterized the neighborhood food availability measure by calculating the numbers of neighborhood supermarkets and convenience stores (separately) within a 5-kilometers Euclidean buffer around the centroid (abbreviated as *within the 5-km buffer* below). I obtained supermarket and convenience store data from the 2010 ReferenceUSA dataset, available at the library of the University of North Carolina at Chapel Hill (Infogroup, 2010). I opted to use ReferenceUSA rather than other food data resource (e.g., Dun & Bradstreet, government food registries) due to its greater accuracy and validity in identifying the type and location of retail food outlet (Fleischhacker et al., 2013, Fleischhacker et al., 2012). I classified the supermarkets and convenience stores according to the six-digit primary Standard Industrial Classification (SIC) code. Some private data companies such as Infogroup and D&B have created their own 2-4 digit extension to the original SIC system as a means to update and expand the system so their customers can more precisely define their business classification. See Table

A2-2b in Appendix 2-2 for the classification of food stores based on the six-digit primary SIC code. I cleaned the longitudinal and latitudinal information of the retail food outlet data to maximize their accuracy (e.g., fix incorrect decimal points of longitudinal and latitudinal data). I calculated the counts of supermarkets and convenience stores within the 5-km buffer using ArcGIS 10.5.

I measured the potential household accessibility to fruits and vegetables in the region by using regional destination accessibility (Ramsey and Bell, 2014). The regional destination accessibility measure in Smart Location Database (SLD) is obtained by calculating the number of employees within a 45-minute travel by automobile (network travel time, decay weighted). This measure of accessibility is based on a network analysis model that considers the attractiveness (number of employees) of each reachable block group and the travel time between each origin block group and all the destination block groups simultaneously. Compared to the traditional measure of the total attractiveness of reachable block groups such as summing up the total number of potential destinations in a certain area, regional destination accessibility calculated by the SLD offers a more accurate measure of total attractiveness by decaying the attractiveness of destinations using distance decay curve. I spatially linked the household's residential zip code to the matching SLD block group to obtain the value of the regional destination accessibility for the household. See Appendix 2-2 for the details of developing regional destination accessibility.

I constructed the availability of neighborhood destinations and neighborhood destination diversity to reflect how the attractiveness of other potential daily or weekly routine destinations in the neighborhood might also affect food purchase. Increasing the number or diversity of other neighborhood destinations may decrease the expenditures on fruits and vegetables by sacrificing the time spent on food purchasing for other activities in the neighborhood, as neighborhood trip destinations may be substitutes (Bernardin Jr et al., 2009). I used five types of daily or weekly routine destinations (i.e., fast food restaurant (Kerr et al., 2012), sit-down restaurant (Kerr et al., 2012), school (DiSantis et al., 2016), child care service (DiSantis et al., 2016), and church (DiSantis et al., 2016)) that people might piece together with grocery shopping in a chained trip to generate the total number of neighborhood destinations and an entropy index of neighborhood destination diversity within the 5-km buffer. The entropy equation was originally applied by other researchers (Cervero, 1988), and has been used in different land use entropy formulations (Ramsey et al., 2014). The formula I used to calculate neighborhood destination diversity is as follows

$$Neighborhood\ destination\ diversity = \frac{-1}{\ln(5)} \sum_{i=1}^{5} p_i . \ln(\, p_i)$$

where p_i is the proportion of the potential destination in one of the five categories (fast food restaurant, sit-down restaurant, school, child care service, church) within the 5-km buffer. Entropy usually ranges in value from zero (total homogeneity, with all land uses in one category) to 1 (maximum heterogeneity, with an even mixture of land use). I obtained restaurant and other non-food destination data from the 2010 ReferenceUSA dataset. I classified the restaurants and other non-food destinations according to their six-digit primary SIC codes. See Table A2-2b in Appendix 2-2 for the classification of restaurants and non-food destinations based on the six-digit primary SIC code. I calculated the count of fast food restaurants, sit-down restaurants, schools, child care services, and churches within the 5-km buffer using ArcGIS 10.5.

In addition, I used the measure of neighborhood street connectivity provided by the SLD to reflect the directness of travelling to destinations and transportation (Frumkin et al., 2004), which may increase the convenience of purchasing fruits and vegetables by decreasing the transportation cost (e.g., time travelled to food stores). The SLD estimated neighborhood street connectivity as the total number of street intersections divided by the total land area at the block group level. I interpolated this street connectivity variable from the block group in which household resided to within the 5-km buffer by averaging the street connectivity for each block group within the 5-km buffer. See Appendix 2-2 for the details of developing neighborhood street connectivity.

2.2.4 Covariates

Household-level variables I used (in keeping with the work of prior researchers) included education level of female head of household (Volpe et al., 2013, Kyureghian et al., 2012), household income (Volpe et al., 2013, Weatherspoon et al., 2013, Kyureghian et al., 2012, Volpe and Okrent, 2012, De Roos et al., 2017), racial identity of household (Weatherspoon et al., 2013, Kyureghian et al., 2012, Volpe and Okrent, 2012), household size (Weatherspoon et al., 2013, Kyureghian et al., 2012), marital status of household head(s) (Kyureghian et al., 2012), presence of children (Lee et al., 2007, De Roos et al., 2017), and number of employees in the household (household head excluded). All the household-level covariates were retrieved from the Nielsen Homescan Consumer Panel Dataset for 2010. See Appendix 2-2 for the details of developing household-level covariates. Neighborhood-level variables, assessed via either the residential census block group or census tract, included the percent of zero-car households from the SLD, and the percent of households below the poverty line from the 2008-2010 American Community Survey. I spatially linked the household's residential zip code to the matching SLD census block group or American Community Survey census tract to obtain the values of percent of zero-car households and percent of households below the poverty line for the household. The area covariate I used was urbanicity in which households resided, which was classified as urbanized area, urban cluster, and non-urban area by the U.S. Census Bureau.

2.2.5 Statistical analyses

I used a separate set of models for fruits and vegetables to predict the self-reported purchase of each. I used multivariate linear mixed effect regression models that estimate the total expenditure on fruits or vegetables in 2010. I modeled the expenditure on fruits or vegetables as a function of availability of neighborhood supermarkets and convenience stores, regional destination accessibility, availability of neighborhood destinations, neighborhood destination diversity, neighborhood street connectivity, and the household-, neighborhood- and area-level covariates (fruits: n=21,824; vegetables: n=21,824). I included random intercepts for each MSA to enable responses to vary with the MSAs in which the households were nested (Feng et al., 2010). I found that distribution of expenditures on fruits and vegetables purchased by household was right-skewed, so I used the logarithmic transformations of expenditures on fruits and vegetables in the final regressions. I opted not to use the Nielsen sampling weight since the initial sample marginal distributions that were used to create the weights were not representative of the sample I retained for my analysis (Kyureghian et al., 2012). All statistical analyses were conducted using Stata, version 14.0. I used the xtmixed package in Stata to run the linear mixed effect regression models.

2.2.6 Sensitivity testing

In addition to the 5-kilometer buffer, I used 3-km buffer to test the sensitivity of the neighborhood definition on the results. In addition, I ran the models on the reduced sample (between 3rd and 97th percentile of the expenditures) to test the effect of outlier definition. As I removed non-magnet households who did not report non-standard UPC products, I examined the similarity of household-level sociodemographic characteristics (i.e., education level of female head of household, household income, race identify of household, household size, marital status of household head(s), presence of children, number of workers in the household) between magnet households who reported non-standard UPC products versus non-magnet households who did not to examine this source of selection bias using kernel density plots of each to compare similarity between these two groups. The overlap of the plotted curves between these two groups reflected the balance of neighborhood-level and individual-level characteristics between these two groups. Additional information about the balance tests can be found in Appendix 2-4.

2.3 Results

2.3.1 Descriptive statistics

The households in the sample was predominantly white and highly educated, 58.3 percent of which had an annual household income greater than 50,000 dollars (see Table 2.1). 59% of the households did not have a supermarket in their neighborhood, yet their expenditures in fruits and vegetables are not substantially different from their peers who have more supermarkets in their neighborhoods (Figure 2.1). Furthermore, households living in urbanized area and non-urban area had 3.1 and 0.6 supermarkets on average within the neighborhood (data not shown). The insignificant difference of expenditures on vegetables between households in urbanized and non-urban areas (t=1.1, p=0.27) also implied that the number of neighborhood supermarkets was only weakly associated with food purchase.

Table 2.1. Sample characteristics

Characteristics	2010
Number of observations	22,448
Annual expenditure on food purchased, $, mean (SD)	
Fruits	133.7 (124.4)
Vegetables	135.8 (118.9)
Availability of neighborhood food stores	
Number of supermarkets within the neighborhood, count, mean (SD)	2.0 (6.8)
0	59.0%
1	18.5%
2+	22.5%
Number of convenience stores within the neighborhood, 10 counts, mean (SD)	1.2 (1.8)
Regional destination accessibility	
Jobs within 45 minutes auto travel time, 10,000 jobs, mean (SD)	9.6 (12.0)

Availability of neighborhood destinations	
Number of fast food restaurants, sit-down restaurants, schools, child care services, and churches within the neighborhood, 10 counts, mean (SD)	21.0 (40.2)
Neighborhood destination diversity	
Entropy within the neighborhood, 10 percent, mean (SD)	7.0 (1.8)
Neighborhood street connectivity	
Street intersection density (weighted, auto-oriented intersections eliminated) within the neighborhood, 10 intersections per square mile, mean (SD)	3.5 (3.0)
Percent of zero-car households, mean (SD)	5.3 (5.3)
Percent of population below poverty level, mean (SD)	8.2 (8.4)
Education level of female head of household [a], %	
≤High school or below	22.9
>High school	68.2
No female head	8.9
Household income, $ [b], %	
Under 20,000	8.0
20,000-49,999	33.7
50,000+	58.3
Race identity of household, %	
White	83.3
Black	9.3
Asian	2.9
Other	4.5
Household size, %	
Single member	21.0
Two members	41.7
Three members	15.7
Four + members	21.7
Marital status of household head(s), %	
Married	64.8
Widowed	5.1
Divorced/separated	15.5
Single	14.5
Presence of children, %	
Yes	26.6
Number of workers in the household (household head excluded), mean (SD)	0.2 (0.5)
Urbanicity, %	
Urbanized area	58.4
Urban cluster	4.2
Non-urban	37.4

Abbreviation: SD: standard deviation.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

Notes [a] For households with two heads of household, Nielsen designates the characteristics of the head of household as whoever makes most of the purchasing decisions.

[b] The value represented ranges of total household income for the full year that is 2 years prior to the panel year.

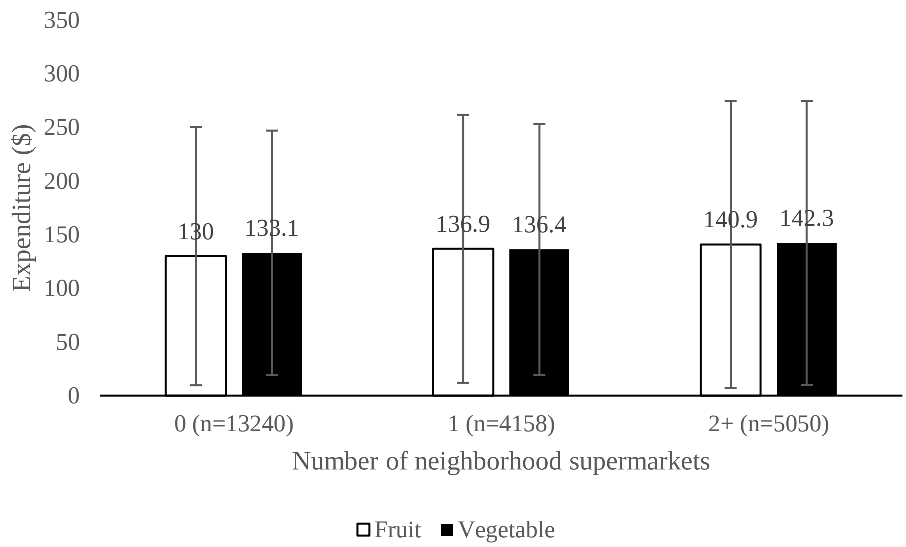

Figure 2.1. Annual expenditures on fruits and vegetables by household by number of supermarkets in the neighborhood

2.3.2 Regression analyses

Tables 2.2 and 2.3 present the results of regression analyses in this study. Analyses including only the availability of neighborhood supermarkets suggested households living in a neighborhood with at least one supermarket purchased significantly more (by 4-5 percentage points) fruits than households living in neighborhoods without supermarket (Table 2.2).

Controlling for the broader built environmental context (i.e., availability of neighborhood convenience stores, regional destination accessibility, availability of neighborhood destinations, neighborhood destination diversity, neighborhood street connectivity, and urbanicity), the difference of expenditures on fruits between households with different numbers of supermarkets was not significant; and I observed that households purchased significantly more (3 percentage points) fruits if they lived in neighborhoods with fewer (10 fewer) convenience stores. I observed that households purchased significantly more (1 percentage point more) fruits if they lived in neighborhoods with greater (10, 000 jobs more) regional destination accessibility. I also observed that households purchased significantly more fruits if they lived in urbanized area than

Table 2.2. Coefficients of cross-sectional associations between annual expenditures on fruits and the presence of supermarkets (5-km buffer), broader built environmental context characteristics, neighborhood- and household-level covariates, by Nielsen household (obs=21,824 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 5-km buffer				
0 (ref)	---	---	---	---
1	0.04 (0.02)	**0.047**	0.02 (0.02)	0.201
2+	0.05 (0.02)	**0.008**	0.03 (0.02)	0.209
Broader built environmental context				
Availability of convenience stores, 10 counts, 5-km buffer			-0.03 (0.01)	**0.000**
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.01 (0.00)	**0.003**
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 5-km buffer			0.00 (0.00)	0.161
Neighborhood destination diversity: Entropy, 10 percent, 5-km buffer			-0.00 (0.00)	0.984
Neighborhood street connectivity: 10 intersections per square mile, 5-km buffer			0.01 (0.00)	**0.002**
Urbanicity				
Urbanized area (ref)			---	---
Urban cluster			-0.08 (0.03)	**0.022**
Non-urban area			-0.00 (0.02)	0.987

Abbreviation: SE: standard error. **Bold** indicates statistical significance ($p<0.05$).
Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.
Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 2nd percentile or greater than the 98th percentile. The 21, 824 households in the fruit model were not necessarily the same 21,824 households in the vegetable model; there were 624 households in the fruit model but not in the vegetable model, and vice versa.
[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

Table 2.3. Coefficients of cross-sectional associations between annual expenditures on vegetables and the presence of supermarkets (5-km buffer), broader built environmental context characteristics, neighborhood- and household-level covariates purchased by Nielsen household (obs=21,824 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 5-km buffer				
0 (ref)	---	---	---	---
1	0.03 (0.02)	0.061	0.03 (0.02)	0.100
2+	0.04 (0.02)	**0.021**	0.03 (0.02)	0.192
Broader built environmental context				
Availability of convenience stores, 10 counts, 5-km buffer			-0.01 (0.01)	0.184
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.00 (0.00)	0.216
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 5-km buffer			-0.00 (0.00)	0.689
Neighborhood destination diversity: Entropy, 10 percent, 5-km buffer			-0.01 (0.00)	**0.016**
Neighborhood street connectivity: 10 intersections per square mile, 5-km buffer			0.01 (0.00)	**0.024**
Urbanicity				
Urban cluster			-0.01 (0.03)	0.690
Non-urban area			0.00 (0.01)	0.937

Abbreviation: SE: standard error. **Bold** indicates statistical significance (p<0.05).
Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.
Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 2nd percentile or greater than the 98th percentile. The 21, 824 households in the fruit model were not necessarily the same 21,824 households in the vegetable model; there were 624 households in the fruit model but not in the vegetable model, and vice versa.
[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

did those living in urban cluster.

Similarly, analyses including only the availability of neighborhood supermarkets suggested households living in a neighborhood with two supermarkets purchased significantly more (4 percentage points more) vegetables than households living in neighborhoods without supermarket (Table 2.3).

Controlling for the broader built environmental context, the difference of expenditures on vegetables between households with different numbers of supermarkets was not significant; and I observed that households purchased significantly fewer (1 percentage point fewer) vegetables if they lived in neighborhoods with lower (10 percentage points of entropy value lower) destination diversity. I also observed that households purchased significantly more (1 percentage point more) vegetables if they lived in neighborhoods with higher (10 intersections more) neighborhood street connectivity.

Tables A2-3a, A2-3b, A2-3c, and A2-3d in Appendix 2-3 present the results of sensitivity analyses, which produced results of similar magnitude, direction, and statistical significance. The major exception was that the expenditure on fruits was inversely associated with the entropy value of neighborhood destination diversity within the 5-km buffer but not within the 3-km buffer.

I also found that the magnet households who reported non-standard UPC products had similar household-level characteristics than those non-magnet households who did not report non-standard UPC products. The plotted curves (showing overlap between these two groups) are shown in Figure A2-4 in Appendix 2-4. This ameliorated concerns regarding off-support inference (Oakes and Johnson, 2006) and allowed me to include all the magnet households into the expenditure on fruits or vegetables models based on the evidence that they did not differ greatly from those non-magnet households.

2.4 Discussion

I found some evidence of cross-sectional associations between availability of neighborhood convenience stores, neighborhood street connectivity, regional destination accessibility, neighborhood destination diversity, and the self-reported food purchases. Broader built environment context of the supermarkets is important to consider when evaluating their relationship to healthy food consumption.

Number of neighborhood supermarkets was not associated with expenditures on fruits or vegetables. My findings add to a small but growing number of cross-sectional studies in the U.S. that have indicated that presence of neighborhood supermarkets is not associated with food purchase or the association is small in magnitude (Lin et al., 2014, Rahkovsky and Snyder, 2015, Kyureghian et al., 2012). Policy interventions that focus exclusively on increasing the presence of supermarkets in a neighborhood therefore should proceed with caution. When controlling for the broader neighborhood characteristics, I do not observe significant association between healthy food expenditures and supermarket availability.

It is possible that households traveled out of the 5 km buffer to purchase food to compensate for the inadequate or unsatisfactory food resources within the buffer. Therefore, they did not necessarily rely on the supermarkets within the 5-km buffer to purchase fruits or vegetables. Kerr et al. , using the food trips of 4,800 residents in the Atlanta region as a case study, observed that the average distance travelled to a grocery store or supermarket was 7.5 kilometers (Kerr et al., 2012). Kerr et al. further found that grocery shopping more often occurs while travelling from locations other than home. Future work should therefore be conducted assessing the availability of supermarkets around work and particularly along frequently travelled routes because households may need to fit food shopping into a plethora of other activities.

I found a positive association between the neighborhood street intersection density and expenditures on fruits and vegetables. Neighborhood street intersection density may facilitate food purchases by shortening the time required to travel to retail food stores through an increased directness of route and number of alternative routes (Thornton et al., 2011). My measure of neighborhood street connectivity was weighted to reflect connectivity for pedestrian and bicycle travel (Ramsey and Bell, 2014); the significant association between neighborhood street connectivity and the expenditure on fruits and vegetables may therefore suggest that walkable streets encouraged people to go out and do grocery shopping even if they do not necessarily travel to food stores by walking or via bicycle (Krizek and Johnson, 2006).

Households who live in a neighborhood with more convenience stores reported purchasing less fruits than did those who lived in neighborhoods with fewer convenience stores. My results were consistent with a previous study, which observed that the neighborhood convenience stores were negatively associated with the self-reported diet quality (Rummo et al., 2015). Future work should be conducted to examine if convenience stores forestalled people's demand for healthy options by offering unhealthy ones, in other words if living in a neighborhood with a greater availability of convenience stores encouraged people to purchase more unhealthy food (e.g., energy dense snacks and sweetened beverages), thus decreasing the purchase of healthy options (e.g., fruits) in other more-distant retail food outlets selling more healthy food options.

Purposeful placement of food stores and restaurants may introduce bias if the factors related to food store and restaurants are directly or indirectly associated with the exposure (e.g., availability of supermarkets and convenience stores) and the outcomes (e.g., expenditures on fruits and vegetables). For example, it is possible that the association I observed between more convenience stores and less purchase of fruits occurred because fresh fruits were less affordable for poor households than wealthy households and because more convenience stores targeted poor households. However, the number of neighborhood convenience stores by household income at the individual level seemed to contradict this possibility. In the sample the average number of neighborhood convenience stores was similar for households with different income levels (5.8, 5.1, and 4.6 convenience stores for those who had household income of under 20,000, 20,000-49,999, and 50,000+, respectively); it was therefore unlikely that low-income households used neighborhood convenience stores (to purchase unhealthy food) more often simply because there were a disproportionally larger number of convenience stores located in their vicinity compared to households with higher income. Future work should be conducted which employs complex models (e.g., instrumental-variable regression, simultaneous equation) to explicitly account for the purposeful placement of food stores and restaurants in neighborhood over time.

However, even if I was able to establish the existence of a robust and inverse association between convenience stores and fruits purchased, limiting the number of convenience stores might not be a desirable policy. Alleviating unhealthy food items without introducing alternative food stores could be problematic, particularly for those lacking access to transportation (Sturm and Cohen, 2009, Fox and Horowitz, 2013). Further, convenience stores in urban areas may have some additional advantages of spatial accessibility, such as the integration of food shopping into other daily activities like meeting children, easy parking, etc. (Cannuscio et al., 2014). Those attempting a policy intervention should therefore consider how to overcome the financial and marketing barriers to increase the provision and promotion of fresh produce in convenience stores as well as change resident's food preference through nutrition education, alteration of social norms, etc. (Ruff et al., 2016, Zenk et al., 2015). Other non-spatially-based interventions could be employed, such as pricing instruments (e.g., beverage tax) (Powell et al., 2013) and education (Martin et al., 2012), to decrease the purchase of unhealthy food in convenience stores.

Increasing regional destination accessibility in the region was associated with greater expenditure on fruits, although the magnitude was small. My results were consistent with those of other researchers, who observed that food opportunities were linked to other places during weekly or daily routine travels beyond home neighborhood (Kerr et al., 2012, Clifton, 2004, DiSantis et al., 2016). My measure of regional destination accessibility used a complex method to specify greater weights for potential destinations closer to the home neighborhood and smaller weights for potential destinations farther from home, which takes into account benefits/costs from both transportation and land-use decisions (Proffitt et al., 2017). Thus, my results suggest that living in a more compact region (i.e., smaller distances to all destinations) provided more destination opportunities, including food stores, which in turn implies that opportunistic food purchases may occur during other activities. Future work should be conducted to identify the non-food purchase opportunities that are related to food purchase opportunities and to refine my measurement of regional destination accessibility, which might increase the explanatory power of regional destination accessibility on food purchasing behaviors.

I did not find that households purchased more fruits or vegetables if they lived in a neighborhood with many other destinations such as fast food and sit-down restaurants, schools, child care services and churches. My results thus suggested that the total number of neighborhood destinations was not associated with the expenditures on fruits or vegetables. However, I found that greater neighborhood destination diversity was associated with lower expenditures on vegetables. It is possible that the large number of other destinations I observed is a cluster of similar resources (e.g., restaurants), but the effects may be different if the households live in a neighborhood with diverse and dissimilar destinations, for example two restaurants, one school, one child care service, and a church, as compared to a neighborhood with five restaurants. Future research should be performed which examines if people living in neighborhood with greater destination diversity tend to eat food in eat-out places (e.g., restaurants, coffee shops) and thus purchase less fruits and vegetables. As previous studies have indicated that the allocation of time had direct consequences for food purchase (Redman, 1980, Nevo and Wong, 2015), time constraints may thus act as an important moderator between the spatial distribution and arrangement of neighborhood destinations, how people move between and spend time on destinations, and the location at which they shop food (Kerr et al., 2012), which has been previously inadequately addressed (Widener and Shannon, 2014).

I did not find an association between neighborhood destination diversity and expenditure on fruits, although I found such an association between neighborhood destination diversity and expenditure on fruits. Nor did I observe an association between number of neighborhood convenience stores and expenditure on vegetables. Nor did I observe an association between regional destination accessibility and expenditures on vegetables. Future research should examine other types of food purchased (e.g., snacks, beverages) and their relationship with the built environment to further my understanding of the relationship between built environment and purchasing behaviors by food type.

Households living in urbanized areas and non-urban areas reported higher expenditures on fruits than those living in urban clusters, suggesting a varied expenditure on fruits within the MSA. My results therefore supported going beyond the basic urban and rural dichotomy used by other researchers and focusing more on the "peri-urban" area which encompass a fragmented mixing of urban and nonurban worlds (Lerner and Eakin, 2011) and the challenge of accessing food sources faced by urban clusters. In addition, as households living in urbanized area and non-urban area had 3.1 and 0.6 supermarkets on average within the 5-km buffer, respectively (data not shown), the insignificant difference of expenditures on fruits and vegetables between urbanized area and non-urban area implied that the number of neighborhood supermarkets was only weakly associated with food purchase. Future work should be conducted to determine which type of food sources (e.g., traditional food store such as supermarket and grocery store, nontraditional food store such as warehouse clubs, supercenters, specialty stores, non-store sources of food such as farmer market) are used by households in different areas.

Overall, my findings suggested that built environmental policy could still be effective but the focus should turn to a comprehensive approach of thinking about food purchase behaviors that incorporates all aspects of built environment in the broader area beyond the home neighborhood and an interconnected context not restricted to the food destination (Institute of Medicine, 2012, Raja et al., 2010). Time constraint may play an important role in linking the built environment and food purchasing behaviors. This systems approach will require urban planners and public health officials to work together to pursue systematic strategies (both spatial and non-spatial) that regulate land use and the transportation systems related to the spatial movement of grocery shoppers, such as retail land use along commuting corridors, the mix of retail and employment in employment neighborhoods (Widener et al., 2013).

The strength of my study comes from the use of a large, nationwide sample, which made possible the construction of household-level measures of the built environment and provided the sample size and the variance in built environment required to measure reliable and small associations between the built environment and the purchase of fruits and vegetables. My built environment data was carefully constructed to reflect the complexity, quality and intensity of environment clues, which helped to identify relationships between built environment factors and food purchasing behaviors.

My study has several limitations. Measurement errors existed in my estimates of food purchase expenditures due to many reasons (e.g., skipping reporting some of the purchases, minor purchases made at convenience stores) (Zhen et al., 2009). This issue is probably more severe in fresh produce (fresh fruits, vegetables, beef, poultry, pork, etc.) than other food products (Zhen et al., 2009). The average annual expenditures on fresh fruits and vegetables using Nielsen data were 32-39 percent lower than the equivalent estimates generated using the Consumer Expenditure Survey (CES) data (U.S. Bureau of Labor Statistics, 2010), which was similar to the discrepancy observed by others (Zhen et al., 2009). I used the centroid of the zip code tabulation area as a proxy for the household's exact residential location, which meant that households living in the same zip code tabulation area shared the same built environment characteristics. My measure of regional destination accessibility may have included destinations unrelated to food purchase. I also could not assess possibly-correlated non-spatial factors influencing food purchases (e.g., food price, purchasing power) which might also be correlated with regional destination accessibility. However, I used income at the household-level, together with poverty level at the neighborhood level to attempt to control for purchasing power. My analytic sample comprised a greater number of middle-income households with a greater average educational attainment than the U.S. national population (Piernas et al., 2013, Ford et al., 2014) and I only selected households who reported magnet data, which limits the generalizability of my results to all U.S. households. I lacked built environment data for years other than 2010, thus I was unable to use longitudinal data to address unmeasured confounders such as residential self-selection. Longitudinal analyses may permit the resolution of the unexpected cross-sectional associations (Zenk et al., 2017), such as the inverse associations between number of neighborhood supermarkets and the expenditures on fruits (although the associations were not significant). Commercial sources of data on food outlets are prone to error; I carefully cleaned and processed the data I used to increase its quality, yet errors likely remained.

2.5 Conclusions

My ability to use food purchase data from a national study lends credence to the theorized association between availability of neighborhood convenience stores, regional destination accessibility, neighborhood destination diversity, neighborhood street connectivity, and purchase of fruits and vegetables. However, I did not find that people living in a neighborhood with many supermarkets reported purchasing more fruits or vegetables. I have added to a small but growing literature demonstrating that interventions that increase the number of neighborhood supermarkets should proceed with caution. Households living in an area with a greater neighborhood street connectivity, regional destination accessibility, and lower availability of neighborhood convenience stores and neighborhood destination diversity may purchase more fruits or vegetables, but the underlying mechanism needs to be examined more thoroughly.

Acknowledgements

The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

REFERENCES

Cannuscio, C. C., Hillier, A., Karpyn, A. & Glanz, K. 2014. The social dynamics of healthy food shopping and store choice in an urban environment. *Social Science & Medicine,* 122**,** 13-20.

Clifton, K. J. 2004. Mobility strategies and food shopping for low-income families: A case study. *Journal of Planning Education and Research,* 23**,** 402-413.

DiSantis, K. I., Hillier, A., Holaday, R. & Kumanyika, S. 2016. Why do you shop there? A mixed methods study mapping household food shopping patterns onto weekly routines of black women. *International Journal of Behavioral Nutrition and Physical Activity,* 13**,** 11.

Ford, C. N., Ng, S. W. & Popkin, B. M. 2014. Are food and beverage purchases in households with preschoolers changing?: A longitudinal analysis from 2000 to 2011. *American journal of preventive medicine,* 47**,** 275-282.

Fox, A. M. & Horowitz, C. R. 2013. Best practices in policy approaches to obesity prevention. *Journal of health care for the poor and underserved,* 24**,** 168.

Institute of Medicine 2012. *Accelerating progress in obesity prevention: solving the weight of the nation,* Washington, DC, The National Academies Press.

Kerr, J., Frank, L., Sallis, J. F., Saelens, B., Glanz, K. & Chapman, J. 2012. Predictors of trips to food destinations. *International Journal of Behavioral Nutrition and Physical Activity,* 9**,** 58.

Krizek, K. J. & Johnson, P. J. 2006. Proximity to trails and retail: effects on urban cycling and walking. *Journal of the American Planning Association,* 72**,** 33-42.

Kyureghian, G., Nayga Jr, R. M. & Bhattacharya, S. 2012. The effect of food store access and income on household purchases of fruits and vegetables: A mixed effects analysis. *Applied Economic Perspectives and Policy,* 35**,** 69-88.

Lerner, A. M. & Eakin, H. 2011. An obsolete dichotomy? Rethinking the rural–urban interface in terms of food security and production in the global south. *The Geographical Journal,* 177**,** 311-320.

Lin, B.-H., Ver Ploeg, M., Kasteridis, P. & Yen, S. T. 2014. The roles of food prices and food access in determining food purchases of low-income households. *Journal of Policy Modeling,* 36**,** 938-952.

Martin, K. S., Havens, E., Boyle, K. E., Matthews, G., Schilling, E. A., Harel, O. & Ferris, A. M. 2012. If you stock it, will they buy it? Healthy food availability and customer purchasing behaviour within corner stores in Hartford, CT, USA. *Public health nutrition,* 15**,** 1973-1978.

Nevo, A. & Wong, A. 2015. The elasticity of substitution between time and market goods: Evidence from the Great Recession. National Bureau of Economic Research.

Oakes, M. & Johnson, P. 2006. Propensity score matching for social epidemiology. Chapter 15 in Methods in Epidemiology, eds. Oakes, MJ and Kaufman, JS New York. John Wiley & Sons, Inc.

Piernas, C., Ng, S. W. & Popkin, B. 2013. Trends in purchases and intake of foods and beverages containing caloric and low-calorie sweeteners over the last decade in the United States. *Pediatric obesity,* 8**,** 294-306.

Powell, L. M., Chriqui, J. F., Khan, T., Wada, R. & Chaloupka, F. J. 2013. Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes. *Obesity reviews,* 14**,** 110-128.

Proffitt, D. G., Bartholomew, K., Ewing, R. & Miller, H. J. 2017. Accessibility planning in American metropolitan areas: Are we there yet? *Urban Studies***,** 0042098017710122.

Rahkovsky, I. & Snyder, S. 2015. Food choices and store proximity. United States Department of Agriculture, Economic Research Service.

Raja, S., Yin, L., Roemmich, J., Ma, C., Epstein, L., Yadav, P. & Ticoalu, A. B. 2010. Food environment, built environment, and women's BMI: evidence from erie county, New york. *Journal of Planning Education and Research*.

Ramsey, K. & Bell, A. 2014. *Smart location database Version 2.0 User Guide* [Online]. Washington, DC: United States Environmental Protection Agency. Available: https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide [Accessed June 15 2017].

Redman, B. J. 1980. The impact of women's time allocation on expenditure for meals away from home and prepared foods. *American Journal of Agricultural Economics,* 62**,** 234-237.

Ruff, R. R., Akhund, A. & Adjoian, T. 2016. Small convenience stores and the local food environment: an analysis of resident shopping behavior using multilevel modeling. *American Journal of Health Promotion,* 30**,** 172-180.

Rummo, P. E., Meyer, K. A., Boone-Heinonen, J., Jacobs Jr, D. R., Kiefe, C. I., Lewis, C. E., Steffen, L. M. & Gordon-Larsen, P. 2015. Neighborhood availability of convenience stores and diet quality: Findings from 20 years of follow-up in the Coronary Artery Risk Development in Young Adults study. *American journal of public health,* 105**,** e65-e73.

Sturm, R. & Cohen, D. A. 2009. Zoning for health? The year-old ban on new fast-food restaurants in South LA. *Health Affairs,* 28**,** w1088-w1097.

Thornton, L. E., Pearce, J. R. & Kavanagh, A. M. 2011. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 71.

U.S. Bureau of Labor Statistics 2010. Consumer Expenditure Survey, 2006-2010 Washington, DC: U.S. Bureau of Labor Statistics.

Widener, M. J., Farber, S., Neutens, T. & Horner, M. W. 2013. Using urban commuting data to calculate a spatiotemporal accessibility measure for food environment studies. *Health & Place,* 21**,** 1-9.

Widener, M. J. & Shannon, J. 2014. When are food deserts? Integrating time into research on food accessibility. *Health & place,* 30**,** 1-3.

Zenk, S. N., Powell, L. M., Isgor, Z., Rimkus, L., Barker, D. C. & Chaloupka, F. J. 2015. Prepared food availability in US food stores: a national study. *American journal of preventive medicine,* 49**,** 553-562.

Zenk, S. N., Tarlov, E., Wing, C., Matthews, S. A., Jones, K., Tong, H. & Powell, L. M. 2017. Geographic Accessibility Of Food Outlets Not Associated With Body Mass Index Change Among Veterans, 2009–14. *Health Affairs,* 36**,** 1433-1442.

Zhen, C., Taylor, J. L., Muth, M. K. & Leibtag, E. 2009. Understanding differences in self-reported expenditures between household scanner data and diary survey data: a comparison of Homescan and consumer expenditure survey. *Review of Agricultural Economics,* 31**,** 470-492.

CHAPTER 3. THE ASSOCIATION BETWEEN NEIGHBORHOOD TYPE AND RELATIVE AVAILABILITY OF SIT-DOWN RESTAURANTS AND SUPERMARKETS IN THE TWIN CITIES REGION OF MINNESOTA

3.1 Introduction

Previous studies on access to healthy food have generally characterized neighborhoods using a single construct of neighborhood context (Lytle, 2009), such as income or race. Although low-income and minority-dominant neighborhoods have been generally identified as having poor access to healthy food (Walker et al., 2010), findings are mixed. For example, several reports have indicated that low-income neighborhoods tend to have a greater availability of fast-food restaurants (Powell et al., 2007, Zenk and Powell, 2008) whereas others do not (Lamichhane et al., 2013, James et al., 2014). In fact, I know little about neighborhoods defined using a more nuanced categorization, which is unfortunate given that neighborhood is patterned across many interrelated built environment and sociodemographic characteristics (Jones and Huh, 2014). The types of restaurants and food stores available in neighborhoods also vary by built environmental factors such as population density and land use pattern in addition to sociodemographic factors. For example, restaurants and food stores may choose to open outlets even in poor neighborhoods, if residential densities are sufficiently high for demand based on larger population (Helling and Sawicki, 2003).

As neither aggregate indices of sociodemographic status (SES) nor specific aspects of the built environment appear in isolation in neighborhoods (Nelson et al., 2006), I used a novel method to classify neighborhood types using a combination of several domains. I used cluster analyses, which has been used by studies (Jones and Huh, 2014, Meyer et al., 2015, Nelson et al.,

2006) as a measurement strategy to disentangle the mixed results of neighborhood effects confounded by correlations among neighborhood features. Cluster analysis takes into account a broad set of neighborhood resource variables to more fully represent the factors (such as population density, mix of land use, and SES) present in a neighborhood on which restaurants and food stores base their decision to open an outlet in that neighborhood and to identify homogenous groups of neighborhoods with shared neighborhood characteristics.

I then looked at the distribution of types of restaurants and food stores within each type of neighborhood to understand which types of neighborhoods had greater access to specific types of restaurants or food stores. I used the Twin Cities Region of Minnesota (abbreviated as Twin Cities Region), an area of nearly three million people living in 186 communities across the seven counties of Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington, as the study area. The Twin Cities Region has developed several distinctive types of neighborhoods (e.g., active downtown, vibrant urban) (Minneapolis Metropolitan Council, 2015). In addition, the neighborhood environment in the Twin Cities Region has become increasingly diverse in social composition and physical form over the past decades (Minneapolis Metropolitan Council, 2015), which provided the opportunity to observe differences of and changes in the distribution of neighborhood food resources over time. Using the data on business types and locations for 1993, 2001 and 2011, I examined the associations between neighborhood types as defined using the 1993 data and the changes in the relative availability of sit-down restaurants and relative availability of supermarkets in the neighborhoods over an 18-year period (1993-2011) in the Twin Cities Region.

3.2 Methods

3.2.1 Study area

My study area included 2,083 census block groups defined in 2010 by the U.S. Census Bureau in the Twin Cities Region, a 7-county (Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington) area with diverse built environment and sociodemographic characteristics (Minneapolis Metropolitan Council, 2015). I used census block groups to operationalize neighborhoods. The census block group (approximate population of 1,500) is the smallest unit for which census data on built environment and sociodemographic measures is available. I excluded only two census block groups in the Twin Cities Region due to missing data.

3.2.2 Relative availability of sit-down restaurants and supermarkets

I obtained food resource data from the Dun & Bradstreet (D&B) Duns Market Identifiers File (restaurant and food store Standard Industrial Classification (SIC) categories; Dun & Bradstreet, Inc., Short Hills, NJ), a secondary commercial data source widely available in the United States. I then classified the food resources according to primary eight-digit SIC codes for data in years 1993, 2001, and 2011. See Appendix Table A1-1 for the SIC codes from D&B.

I sought to characterize the neighborhood restaurant and food store availability by calculating the relative availability of sit-down restaurants (relative to total restaurants) and supermarkets (relative to total food stores). Recent reports suggest that relative measures, such as ratio or proportion of various types of food retail outlets, may be as or more important to diet-related behaviors than the number of outlets because they offer neighborhood residents competing options (Clary et al., 2015, Mercille et al., 2012, Rummo et al., 2015). I defined the relative availability of sit-down restaurants as the percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants in the neighborhood (abbreviated as percentage of sit-down restaurants below). I defined the relative availability of supermarkets as the percentage of supermarkets relative to total supermarkets, grocery stores, and convenience stores in the neighborhood (abbreviated as percentage of supermarkets below). I calculated the count of each type of food resource within each neighborhood in each observational year using ArcGIS 10.3, then used the counts to calculate the percentages of sit-down restaurants and supermarkets.

3.2.3 Neighborhood type

I classified neighborhood type using a cluster analysis that included 13 built environment and sociodemographic characteristics in 1990.

3.2.3.1 Neighborhood built environmental characteristics

Neighborhood built environment characteristics included residential population density, employment population density, mix of land use, and percent of single-family housing in the neighborhood. I measured residential population density as the total residential population divided by the total land area of the block group and I measured employment population density as the total employed civilian labor force 16 years and above divided by the total land area of the block group. I retrieved the census population and the size of land area data for census year 1990 from the Longitudinal Tract Database, which allowed me to normalize the census data in 1990 to the boundaries of census tracts in 2010. I then interpolated the normalized census population density data from the census tract level to the census block-group level. I measured the mix of land use using the 3-tier land use entropy equation (with the denominator set to the static 3 land use types in the block group), which used three land use categories (residential, employment and retail) to calculate mix of land use in the block group. The entropy equation was originally applied by Robert Cervero (Ramsey and Bell, 2014), and has been used in different land use entropy formulations. Land use entropy ranges in value from zero (total homogeneity, with all land use in one category) to 1 (maximum heterogeneity, with an even mixture of land use). I measured the percent of single-family housing using the number of single-family housing units divided by the total number of single-family and multi-family housing units. I obtained the data on the category and area of different types of land uses that I used to create the mix of land use and percent of single-family housing from the GIS-based current land-use map in 1990 from the Minneapolis Metropolitan Council.

3.2.3.2 Neighborhood sociodemographic characteristics

Neighborhood sociodemographic characteristics included percent of population aged under 14, aged between 15 and 29, aged between 30 and 44, aged between 45 and 64, aged 65 or above, percent of education of college or above, percent of white race, percent of black race, and median household income. I retrieved all the census sociodemographic characteristics in 1990 from the Longitudinal Tract Database. I then interpolated the normalized census sociodemographic characteristics data from the census tract level to the census block-group level.

3.2.3.3 Cluster analyses

I first transformed each built environment and sociodemographic variable into a z-score to achieve more comparable scales and ranges; otherwise, variables with large ranges might have weighed heavier in the analysis than those with small ranges. I then used the transformed data to conduct partition cluster analyses (the most commonly used form of a method for combining neighborhoods into groups based on their similarity) within the 13 built environment and sociodemographic characteristics, using K-means in Stata. I tested a range of number of clusters, from four to seven, and found that the best clustering solution was a six-cluster solution based on the interpretability of the results and the associated cluster statistics.

3.2.4 Covariates

Neighborhood type in 1993, while robust, did not gradually change over time and therefore could not be used to explain the change in percentages of sit-down restaurants and supermarkets between 1993 and 2011. I therefore added variables for the changes in residential population density, median household income, percent of white race, and percent of single-family housing from 1990 to my models to represent the changes in neighborhood characteristics between 1993 and 2011. I obtained the data on residential population densities from Census 2000 and the 2006−2009 American Community Survey of the U.S. Census Bureau. I calculated the residential population density in 2000 and 2006-2009 in the same way as what I did for residential population density in 1990. I calculated the changes in residential population density in 1990, 2000, and 2006-2009 as zero, the residential population density in 2000 minus the residential population density in 1990, and the residential population density in 2006-2009 minus the residential population density in 1990, respectively. I used the same data source and method to calculate changes in median household income and percent of white race as what I did in calculating changes in residential population density for each neighborhood. I used GIS-based current land-use maps in 2000 and 2010 from the Minneapolis Metropolitan Council to generate percent of single-family housing in these two years; I calculated the changes in percent of single-family housing for each neighborhood in the same way as what I did in calculating changes in residential population density, median household income, and percent of white race.

I also added the total number of sit-down restaurants and fast food restaurants as one of the covariates in the sit-down restaurant model because using only the percentage of sit-down restaurants did not indicate the total number of sit-down restaurants and fast food restaurants (since the percentage was a dimensionless number representing only the proportion sit-down restaurants constituted of the total number of sit-down restaurants and fast-food restaurants).

Similarly, I added the total number of supermarkets, grocery stores and convenience stores as one of the covariates in the supermarket model because the percentage of supermarkets could not adequately indicate the total number of supermarkets, grocery stores and convenience stores.

3.2.5 Statistical analyses

All descriptive analyses and multivariable models were performed using Stata 12.0 (StataCorp, College Station, TX).

3.2.5.1 Descriptive statistics

I generated descriptive statistics to assess neighborhood characteristics in 1990/1993, 2001 and 2011 for all neighborhoods. I calculated means and standard deviations (for continuous variables) of neighborhood built environment characteristics, neighborhood sociodemographic characteristics, and the relative availability of sit-down restaurants and supermarkets in the neighborhood. I used two-tailed Student's t-test to test for statistically significant differences for neighborhood built environment characteristics, sociodemographic characteristics, and the relative availability of sit-down restaurants and supermarkets in the neighborhood between the first (1990/1993) and the last (2011) observational years.

3.2.5.2 Relationship between neighborhood type and relative availability of sit-down restaurants and supermarkets

I used multivariable linear mixed effects regression models to estimate the associations between neighborhood type in 1993 and the percentage of sit-down restaurants and percentage of supermarkets in 1993, 2001, and 2011 (n=2,083). These models appropriately accounted for the clustered data structure of repeated measurements (three measures of food resource data, 1993, 2001 and 2011) for each neighborhood. I modeled the percentage of sit-down restaurants/supermarkets in each neighborhood as a function of neighborhood type in 1993, the time elapsed in years from 1993, the term for the interaction of neighborhood type in 1993 with time elapsed, and the time-varying covariates. I included random intercepts for each neighborhood in the sit-down restaurant and supermarket models to enable the responses to vary within neighborhoods. The model results of the relative availability of sit-down restaurants and supermarkets are presented as (1) A table of the post-estimated linear contrasts of percentage of sit-down restaurants and percentage of supermarkets in the neighborhood by year and for each neighborhood type pair from the multivariable linear mixed effects regression models (Table 3.2); (2) Figures of the estimated mean of percentages of sit-down restaurants and supermarkets over time for each neighborhood type from the same models (Figure 3.2 and 3.3); (3) Tables of the coefficients of the same models using inner city as the reference neighborhood type (Table A3-2a and A3-2b in Appendix 3-2); (4) A table of the p values for the changes of differences in estimated mean of percentage of sit-down restaurants/supermarkets for each neighborhood type pair between two observation years from the same models (Table A3-3 in Appendix 3-3).

3.3 Results

3.3.1 Descriptive statistics

The percentages of sit-down restaurants and supermarkets in the study area increased 10.1 and 3.3 percentage points in 2011, respectively, compared to 1993 (Table 3.1). My study area's population in 2011 (compared to 1993) tended to be older (45-64 or 65 or above), non-white, college educated or higher, possessed of a higher household income; the study area had a

Table 3.1. Selected characteristics of neighborhoods in years 1993, 2001 and 2011, Twin Cities Region

Neighborhood characteristic	**1993 [a]**	**2001**	**2011**	**Change [b]**	**P value [c]**
Number of observations (neighborhoods)	2,083	2,083	2,083	0	---
Relative availability of sit-down restaurants and supermarkets					
Percentage of sit-down restaurants [d], mean (SD)	16.1 ± 33.1	22.7 ± 36.4	26.2 ± 36.8	10.1 ± 41.4	<0.05
Percentage of supermarkets [e], mean (SD)	2.0 ± 11.9	2.4 ± 12.9	5.3 ± 19.6	3.3 ± 20.1	<0.05
Built environment characteristics					
Residential population density, 1,000 person/km^2, mean (SD)	1.5 ± 1.3	1.6 ± 1.4	1.6 ± 1.4	0.1 ± 0.4	<0.05
Employment population density, 1,000 person/km^2, mean (SD)	0.8 ± 0.8	0.8 ± 0.8	0.9 ± 0.8	0.1 ± 0.3	<0.05
Mix of land use [f], mean (SD)	44.2 ± 27.9	48.0 ± 29.7	53.0 ± 28.2	8.8 ± 20.4	<0.05
Percent of single-family housing [g], mean (SD)	61.3 ± 33.8	63.7 ± 33.6	77.3 ± 31.5	16.3 ± 27.0	<0.05
Total sit-down restaurants and fast food restaurants, mean (SD)	0.8 ± 2.1	1.3 ± 2.6	2.1 ± 4.5	1.3 ± 3.2	<0.05
Total supermarkets, grocery stores and convenience stores, mean (SD)	0.8 ± 1.1	1.0 ± 1.3	1.0 ± 1.3	0.2 ± 1.2	<0.05
Sociodemographic characteristics					
Age, mean (SD)					
Percent of population under 14	22.2 ± 6.4	21.3 ± 6.8	19.9 ± 6.0	-2.3 ± 4.8	<0.05
Percent of population 15−29	23.7 ± 7.3	20.9 ± 7.8	21.6 ± 8.4	-2.1 ± 4.6	<0.05
Percent of population 30−44	27.2 ± 4.7	26.0 ± 4.8	21.9 ± 4.5	-5.3 ± 5.0	<0.05
Percent of population 45−64	16.9 ± 5.0	21.2 ± 5.4	25.8 ± 5.7	8.9 ± 5.8	<0.05
Percent of population 65 or above	9.9 ± 6.6	10.5 ± 7.3	10.8 ± 6.0	0.9 ± 5.7	<0.05
Percent of population with education level of college or above, mean (SD)	57.6 ± 15.2	66.1 ± 14.9	68.1 ± 14.3	10.5 ± 8.6	<0.05
Race, mean (SD)					
Percent of white race	91.6 ± 13.0	84.1 ± 17.1	80.8 ± 17.6	-10.8 ± 11.6	<0.05
Percent of black race	4.0 ± 9.1	6.1 ± 9.5	8.1 ± 11.1	4.1 ± 7.2	<0.05
Median household income [h], $1,000, median (SD)	38.2 ± 12.5	40.5 ± 15.1	37.1 ± 14.5	0.4 ± 7.5	<0.05
Time elapsed from 1993, year, mean (SD)	0 ± 0	8 ± 0	18 ± 0	18 ± 0	---

Notes [a] I assumed that the neighborhood built environment and sociodemographic characteristics in 1993 were the same as those in 1990.
[b] Change in neighborhood characteristics from year 1993 to 2011.
[c] P value for two-tailed Student's t-test of difference from years 1993 and 2011.
[d] Percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants.
[e] Percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores.
[f] The mix of land use was measured by 3-tier land use entropy (denominator set to the static 3 land use type in the census block group), which used three land use categories (residential, employment and retail) to calculate mix of land use in the block group.
[g] Percent of single-family housing relative to total single-family and multi-family housings.
[h] The median household income in 1993 and 2001 were adjusted for inflation to compare with that in 2011.

greater population density, greater mix of land use, and greater percent of single-family housing in 2011 compared to 1993.

3.3.2 Results from cluster analyses: neighborhood type (Year 1993)

The six robust neighborhood types defined by the final cluster solution represented non-overlapping groupings of Twin Cities Region neighborhoods based on built environment and sociodemographic attributes in 1990 (the first observational year). These clusters included: cluster 1 - high-density urban core; cluster 2 - low-SES inner city; cluster 3 - urban; cluster 4 - aging suburb; cluster 5 - high-income suburb; and cluster 6 - suburban edge. I assumed that the neighborhood type classified by using neighborhood built environment and sociodemographic data in 1990 equaled to neighborhood type classified by using the data in 1993.

I labeled clusters based on their most prominent built environment and sociodemographic characteristics. Cluster 1, which I labeled “high-density urban core”, had greater levels of residential and employment population densities, a greater degree of mix of land use, less single-family housing, more population aged between 15 and 29, and fewer population aged under 14 than most of the other clusters. Cluster 2, labeled as “low-SES inner city”, had moderate-to-high residential and employment population densities and more non-white race population than other clusters (Table A3-1 in Appendix 3-1). Cluster 5 and Cluster 6, which I labeled “high-income suburb” and “suburban edge”, had relatively low levels of residential and employment population densities, lesser degree of mix of land use, and greater median household income than other four types of neighborhoods. Cluster 3 (“urban”) and Cluster 4 (“aging suburb”) had moderate levels of almost all neighborhood features, except for a greater mix of land use and more residents aged 65 or above than other clusters, respectively.

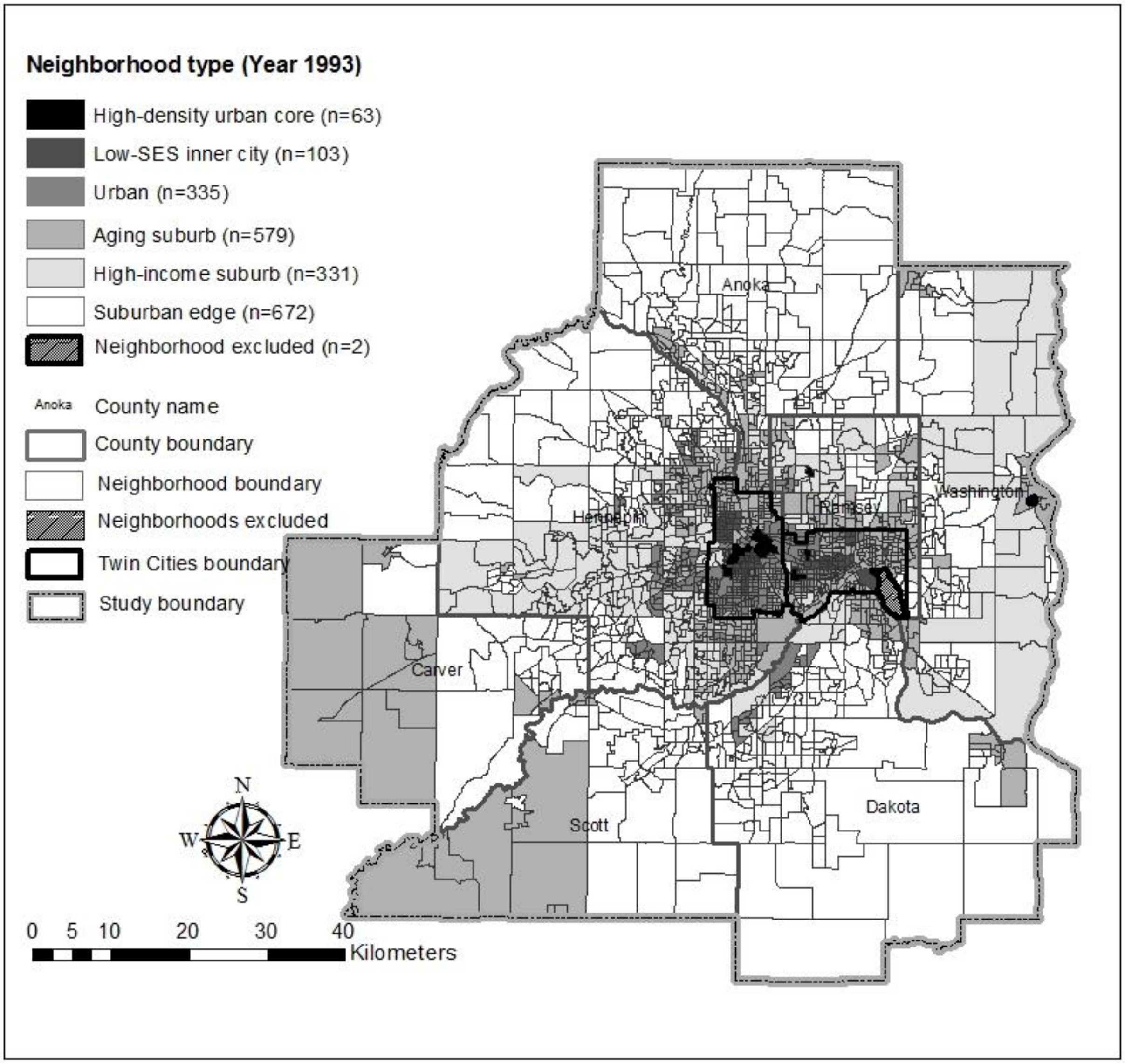

Figure 3.1. Neighborhood types in 1993 in the Twin Cities Region of Minnesota

Figure 3.1 shows a clear geospatial pattern. The high-density urban core (abbreviated as urban core below) and low-SES inner city (abbreviated as inner city below) neighborhoods were tightly clustered in a small segment within the municipal boundaries of the Twin Cities. Urban and aging suburb neighborhoods comprised those transitional areas located between the urban core or inner city neighborhoods and the suburban areas. Another small grouping of aging suburb and high-income extended into the counties of Carver and Scott and the county of

Washington, respectively. The generated clusters reflected comprehensive but distinguishably-different physical and sociodemographic environments.

3.3.3 Relationship between neighborhood type (1993) and relative availability of sit-down restaurants and supermarkets

After controlling for the time-varying variables, urban core neighborhoods had a greater (22.78-27.95 percentage points) percentage of sit-down restaurants than other five types of neighborhoods in 1993; I did not observe any differences in the percentage of supermarkets in 1993. In 2001, I observed more differences in percentages of sit-down restaurants and supermarkets by neighborhood type—inner city neighborhoods had a greater (8.19 percentage points) percentage of sit-down restaurants than suburban edge neighborhoods compared to 1993 (Table 3.2); aging suburb neighborhoods had slightly more (1.52-1.78 percentage points) supermarkets than did the urban, high-income suburb, and suburban edge neighborhoods. In 2011, inner city neighborhoods had more (8.57-12.27 percentage points) sit-down restaurants than did the urban, aging suburb, and high-income suburb neighborhoods. Although urban core neighborhoods had a consistently greater percentage of sit-down restaurants than other neighborhoods in all (three) observational years, the differences between urban core and the other three types (inner city, high-income suburb, and suburban edge) of neighborhoods decreased in 2011 compared to 1993 and 2001. I can see such decreased differences by examining Figure 3.2; as shown in Table A3-3 in Appendix 3-3, such decreased differences were significant.

Table 3.2. Post-estimated [a] linear contrasts of percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants and percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores in the neighborhood by year and for each neighborhood type pair [b] in 1993: Twin Cities Region, 1993-2011

	Sit-down restaurants [c]		**Supermarkets [d]**	
	Estimated beta (95% confidence interval)	P value	Estimated beta (95% confidence interval)	P value
1993				
Urban core vs. inner city	23.02 (13.18, 32.85)	**0.000**	1.87 (-2.50, 6.24)	0.401
Urban core vs. urban	22.78 (13.97, 31.60)	**0.000**	1.41 (-2.51, 5.33)	0.480
Urban core vs. aging suburb	23.75 (14.91, 32.59)	**0.000**	0.18 (-3.75, 4.11)	0.927
Urban core vs. high-income suburb	24.93 (15.47, 34.39)	**0.000**	0.58 (-3.63, 4.78)	0.788
Urban core vs. suburban edge	27.95 (18.61, 37.29)	**0.000**	1.68 (-2.46, 5.83)	0.426
Inner city vs. urban	-0.23 (-7.27, 6.80)	0.948	-0.46 (-3.59, 2.67)	0.773
Inner city vs. aging suburb	0.73 (-6.19, 7.66)	0.835	-1.69 (-4.77, 1.39)	0.282
Inner city vs. high-income suburb	1.91 (-5.65, 9.48)	0.620	-1.30 (-4.66, 2.07)	0.450
Inner city vs. suburban edge	4.94 (-2.41, 12.28)	0.188	-0.19 (-3.46, 3.08)	0.909
Urban vs. aging suburb	0.967 (-3.29, 5.23)	0.656	-1.23 (-3.12, 0.66)	0.203
Urban vs. high-income suburb	2.15 (-2.85, 7.14)	0.400	-0.84 (-3.06, 1.38)	0.461
Urban vs. suburban edge	5.17 (0.63, 9.71)	**0.026**	0.27 (-1.75, 2.29)	0.793
Aging suburb vs. high-income suburb	1.18 (-3.11, 5.47)	0.590	0.39 (-1.51, 2.30)	0.685
Aging suburb vs. suburban edge	4.20 (0.55, 7.86)	**0.024**	1.50 (-0.13, 3.13)	0.071
High-income suburb vs. suburban edge	3.02 (-1.10, 7.14)	0.150	1.11 (-0.73, 2.94)	0.237
2001				
Urban core vs. inner city	13.94 (6.04, 21.84)	**0.001**	2.02 (-1.48, 5.51)	0.258
Urban core vs. urban	18.14 (10.80, 25.48)	**0.000**	2.09 (-1.15, 5.34)	0.206
Urban core vs. aging suburb	19.51 (11.96, 27.06)	**0.000**	0.32 (-3.03, 3.66)	0.853
Urban core vs. high-income suburb	18.81 (10.65, 26.96)	**0.000**	1.91 (-1.70, 5.52)	0.300
Urban core vs. suburban edge	22.13 (14.04, 30.22)	**0.000**	1.84 (-1.74, 5.41)	0.314
Inner city vs. urban	4.20 (-1.44, 9.85)	0.145	0.78 (-2.42, 2.58)	0.951
Inner city vs. aging suburb	5.57 (-0.11, 11.26)	0.055	-1.70 (-4.22, 0.82)	0.186
Inner city vs. high-income suburb	4.87 (-1.38, 11.13)	0.127	-0.11 (-2.88, 2.66)	0.940
Inner city vs. suburban edge	8.19 (2.10, 14.28)	**0.008**	-0.18 (-2.88, 2.52)	0.898
Urban vs. aging suburb	1.37 (-2.06, 4.79)	0.433	-1.78 (-3.29, -0.26)	**0.021**
Urban vs. high-income suburb	0.67 (-3.44, 4.78)	0.750	-0.19 (-2.00, 1.63)	0.842
Urban vs. suburban edge	3.99 (0.24, 7.74)	**0.037**	-0.25 (-1.91, 1.40)	0.763
Aging suburb vs. high-income suburb	-0.70 (-4.18, 2.78)	0.693	1.59 (0.05, 3.13)	**0.043**
Aging suburb vs. suburban edge	2.62 (-0.34, 5.58)	0.082	1.52 (0.21, 2.84)	**0.023**
High-income suburb vs. suburban edge	3.32 (0.03, 6.61)	**0.048**	-0.07 (-1.53, 1.39)	0.925
2011				
Urban core vs. inner city	2.59 (-7.69, 12.86)	0.622	2.19 (-2.37, 6.76)	0.346
Urban core vs. urban	12.33 (2.93, 21.74)	**0.010**	2.95 (-1.24, 7.13)	0.167
Urban core vs. aging suburb	14.21 (4.67, 23.74)	**0.004**	0.48 (-3.76, 4.72)	0.824
Urban core vs. high-income suburb	11.15 (1.09, 21.21)	**0.030**	3.57 (-0.89, 8.04)	0.117
Urban core vs. suburban edge	14.85 (5.04, 24.67)	**0.003**	2.03 (-2.32, 6.39)	0.360
Inner city vs. urban	9.75 (2.49, 17.01)	**0.008**	0.75 (-2.48, 3.98)	0.649
Inner city vs. aging suburb	11.62 (4.46, 18.79)	**0.001**	-1.71 (-4.90, 1.47)	0.292
Inner city vs. high-income suburb	8.57 (0.87, 16.26)	**0.029**	1.38 (-2.04, 4.80)	0.430
Inner city vs. suburban edge	12.27 (4.95, 19.57)	**0.001**	-0.16 (-3.42, 3.10)	0.923
Urban vs. aging suburb	1.87 (-2.56, 6.30)	0.407	-2.46 (-4.43, -0.50)	**0.014**
Urban vs. high-income suburb	-1.18 (-6.47, 4.11)	0.662	0.63 (-1.72, 2.98)	0.600
Urban vs. suburban edge	2.52 (-2.14, 7.17)	0.289	-0.91 (-2.98, 1.16)	0.388
Aging suburb vs. high-income suburb	-3.05 (-7.65, 1.54)	0.193	3.09 (1.05, 5.14)	**0.003**
Aging suburb vs. suburban edge	0.65 (-3.18, 4.47)	0.741	1.55 (-0.15, 3.26)	0.074
High-income suburb vs. suburban edge	3.70 (-0.60, 8.00)	0.092	-1.54 (-3.46, 0.38)	0.115

Notes [a] Multivariable linear mixed effects regressions modeling the percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants and percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores as functions of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white race and percent of single-family housing since 1993, total sit-down restaurants and fast food restaurants (sit-down restaurant model only), and total supermarkets, grocery stores and convenience stores (supermarket model only) and a random intercept for each neighborhood.

[b] Derived from cluster analysis of block-group level data from 1993: percent of population aged under 14, aged between 15 and 29, aged between 30 and 44, aged between 45 and 64, aged above 65, percent of education of college or above, percent of white race, percent of black race, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.

[c] Percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants in the neighborhood.

[d] Percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores in the neighborhood.

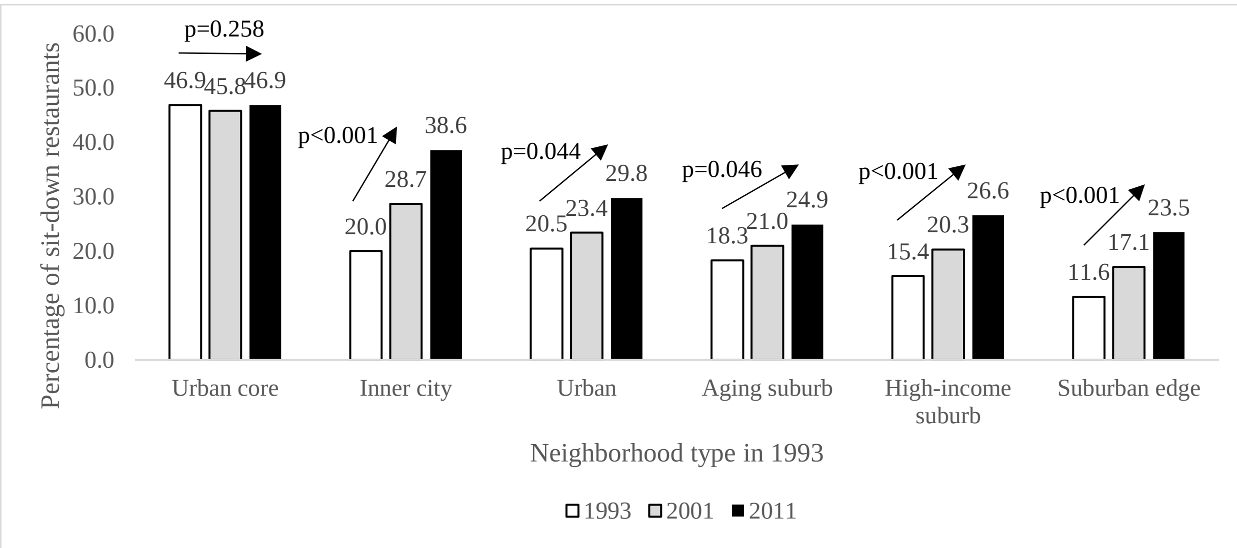

Figure 3.2. Estimated mean [a] of percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants by six types of neighborhoods [b]: Twin Cities Region, 1993-2011.

Notes [a] Multivariable mixed effects regression modeling percentage of sit-down restaurants in each neighborhood as a function of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white and percent of single-family housing since 1993, total sit-down restaurants and fast food restaurants and a random intercept for each neighborhood.

[b] Derived from cluster analysis of block-group level data in 1993: percent of age under 14, age between 15 and 29, age between 30 and 44, age between 45 and 64, age 65 or above, percent of education of college and above, percent of white, percent of black, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.

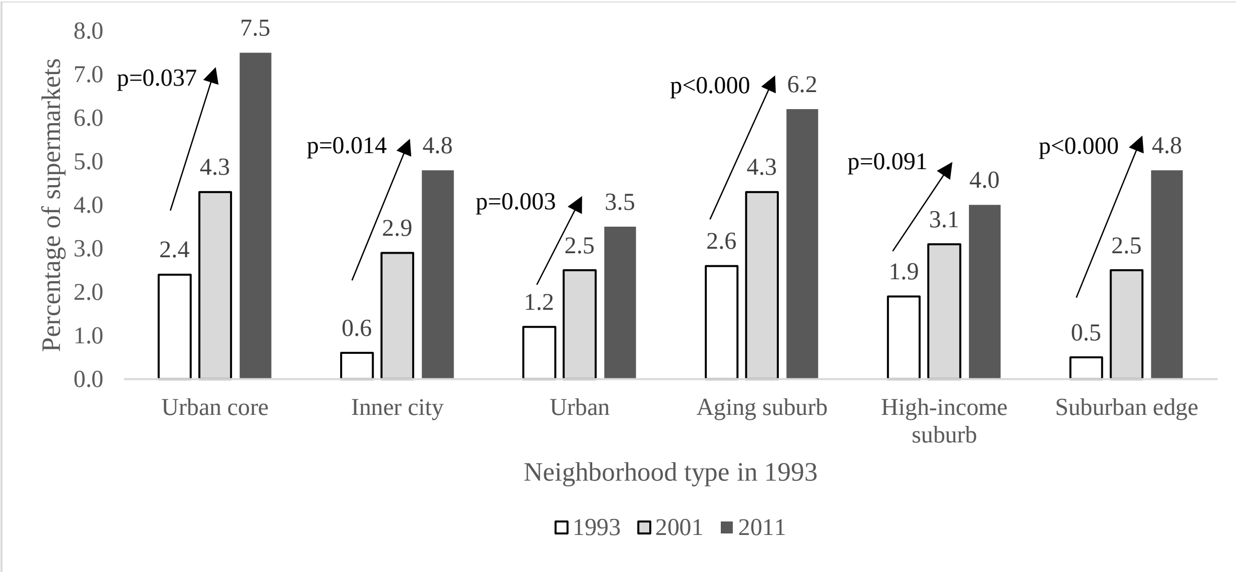

Figure 3.3. Estimated mean [a] of percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores by six types of neighborhoods [b]: Twin Cities Region, 1993-2011.

Notes [a] Multivariable mixed effects regression modeling percentage of supermarkets in each neighborhood as a function of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white and percent of single-family housing since 1993, total supermarkets, grocery stores and convenience stores and a random intercept for each neighborhood.

[b] Derived from cluster analysis of block-group level data in 1993: percent of age under 14, age between 15 and 29, age between 30 and 44, age between 45 and 64, age 65 or above, percent of education of college or above, percent of white, percent of black, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.

3.4 Discussion

I identified six types of neighborhoods in the Twin Cities Region of Minnesota which were characterized by clusters reflecting distinct combinations of built environment and sociodemographic features. Although some of my results indicate an increasingly varied distribution of restaurants and food stores by neighborhood type over time, they also hint at the complexity of the co-varying relationship between the neighborhood built environment and sociodemographic characteristics and the presence of a certain type of food outlet in the neighborhood.

My findings contribute to a growing literature on the associations between the multifaceted composition of the built environment, sociodemographic features and the distribution of food resources. In previous work, researchers investigating the association between neighborhood characteristics and neighborhood food availability have generally characterized the neighborhood features more narrowly, focusing on a single neighborhood construct such as income or race (Lytle, 2009). These studies have produced mixed results. Recognizing that analyses may be confounded by correlations among neighborhood features, I included a broad set of neighborhood resource variables to more fully represent neighborhood-defining patterns based on many interrelated built environment and sociodemographic characteristics.

I found that my neighborhood types were not spatially clustered into homogeneous regions but instead were more distributed across the Twin Cities Region, i.e., the municipal boundaries of the Twin-Cities did not only contain urban core and inner city neighborhoods but also included urban and aging suburbs; similarly, aging suburb and high-income neighborhoods extended to the boundaries of the region, jumping over some of the suburban edge neighborhoods which were least densely populated. This suggests that jurisdictional boundaries of cities are not appropriate to delineate the urban and suburban distinctions in neighborhoods.

I found a more varied distribution of restaurants across neighborhoods in 2001 and 2011 not present in 1993, suggesting some neighborhoods became relatively more appealing to sit-down restaurants and perhaps less appealing to fast food restaurants over time. Specifically, I found only suburban edge neighborhoods had a lower percentage of sit-down restaurants than did the inner city neighborhoods in 2001; however, the urban, aging suburb, and high-income suburb neighborhoods, similar to the suburban edge, also had a lower percentage of sit-down restaurants than did the inner city in 2011. This is because the inner city experienced a greater increase in percentage of sit-down restaurants than did the urban core, urban, and aging suburb neighborhoods between 1993 and 2011 (Table A3-2a in Appendix 3-2). During the period of time of my study, extensive gentrification encouraged new urban development and the improvement of infrastructure (particularly light rail, the park systems and new sports stadium) inside the Twin Cities (Maciag, 2015). Thus, it is possible that infrastructure investments promoted the transformation of the built environment and social composition in inner city neighborhoods, which, in turn, possibly influenced sit-down restaurants to locate in such neighborhoods.

I also found a more varied distribution of food stores across neighborhoods in 2001 and 2011 not present in 1993. Specifically, I found aging suburb neighborhoods had a greater percentage of supermarkets than did the urban and high-income suburb neighborhoods in 2001 and 2011, which I failed to observe in 1993. I found that such differences were largely driven by the great increase of the number of grocery stores and convenience stores in the high-income and suburban edge neighborhoods rather than the great increase of the number of supermarkets in aging suburb (data not shown). Thus, although more grocery stores and convenience stores were situated in aging suburb neighborhoods over time, even more grocery stores and convenience stores were situated in high-income suburb and suburban edge neighborhoods in the same time period.

The percentage of sit-down restaurants in urban core neighborhoods was stable during the observational period; in contrast, other neighborhoods experienced significant increases in the percentage of sit-down restaurants, which in turn decreased the differences in percentage of sit-down restaurants between urban core and other neighborhoods over time. None of the time-varying covariates (i.e., changes in residential population density, median household income, percent of white and percent of single-family housing units) was associated with the change in percentage of sit-down restaurants over time. For example, although the percentage of sit-down restaurants in urban core changed little, residential population densities in urban core increased by between 12.6% and 15.0% in the years between 1993 and 2011; in contrast, residential population densities in the high-income suburb increased little between 1993 and 2011, unlike the significant increase in the percentage of sit-down restaurants in high-income suburb in the same period (data not shown). Thus, decisions by increasingly more sit-down restaurants rather than fast food restaurants to be situated in all neighborhoods except for urban core is still unclear and needs to be examined in future research.

In the supermarket model, however, I found that an increased percentage of supermarkets was associated with a smaller increase (or more rarely a decrease) in the percentage of single-family housing units (Table A3-2b in Appendix 3-2). These less compatible land uses—single-family housing and supermarkets—may have opened up opportunities for urban planners to use policy tools (e.g., zoning) to introduce targeted food stores into the neighborhoods and avoid the difficulty of introducing a supermarket into the neighborhoods with a large increase of single-family housing due to concerns/requirements such as intrusive light (Harvey, 2017), sufficient parking (Cameron et al., 2010), or increased traffic.

Although I did not intend to examine the association between the individual neighborhood characteristics and relative availability of sit-down restaurants and supermarkets, I noticed that some individual neighborhood characteristics may co-vary with each other and jointly affect the distribution of food resources. For example, the urban core had the greatest residential and employment population densities, as well as the greatest percent of population aged between 15 and 29 years, which may jointly contribute to the fact that the highest percentage of sit-down restaurants was in the urban core. Future examination of these associations should use individual-level data that targets restaurant users to disentangle the complexity undergirding the relationships.

Even though the Twin Cities Region experienced multiple economic conditions during the period of my study—the period between 1993 and 2007 in the Twin Cities Region was a time of economic expansion, the period between 2007 and 2009 was a time of economic recession, and the period between 2009 and 2011 was a time of economic recovery (Minnesota Department of Employment and Economic Development, 2014)— I still observed a consistent increase in numbers of sit-down restaurants, fast food restaurants, supermarkets, grocery stores and convenience stores across all neighborhood types (data not shown). This was consistent with national reports and reflected the macroeconomic shifts in the retail food industry (U.S. Department of Agriculture, 2017). Thus, neighborhoods had an increasingly easy access to all foods regardless of neighborhood type over time (Ploeg et al., 2009).

My study had several limitations. First, I only assessed one individual region (Twin Cities Region), which limits comparisons with other geographic areas. The Twin Cities Region had notably more affordable costs for housing and transportation as well as more diverse housing choices than other comparable metropolitan areas (Minneapolis Metropolitan Council, 2015), which may foster a more convenient access to restaurants and small food stores. Second, my data-driven approach to deriving multi-variate groupings may not generalize to other populations. Indeed, my class structure is difficult to compare to those based on single features used by previous researchers. However, given the lack of consistent association in the literature between individual neighborhood characteristics and specific food resource types (Gustafson et al., 2012), I elected to use my data-driven approach to characterize the neighborhood environment. Another major concern is the marked undercount of food outlets in the D&B data, which has the potential to introduce bias (Liese et al., 2010). I used the relative number (expressed as a percentage) of sit-down restaurants and supermarkets to determine if different neighborhood types had different relative numbers of these food resources. If sit-down restaurants had a higher matched rate than fast food restaurants in urban core neighborhoods than in high-income suburb neighborhoods in the D&B data, for example, I risked exaggerating the gap in the numbers of sit-down restaurants relative to total sit-down restaurants and fast food restaurants between urban core and high-income suburbs. Indeed, Powell and colleagues validated the D&B food resource data using direct field observation in the Chicago Metropolitan Statistical Area and found that the matched rate of fast food restaurants differed by various neighborhood characteristics such as income, race and location (urbanized area, urban cluster and non-urban area as defined by the US Census Bureau) (Powell et al., 2011). Since I used 13 built environment and sociodemographic characteristics to classify neighborhoods, future researchers should explore whether the food outlet numbers vary by the overall characteristics of the neighborhood. Fourth, the block group is probably too small to reflect the service area of restaurants and food stores, especially in suburban areas, but using census block group level data yields a better estimate of where food resources and household are located (Ver Ploeg, 2010), compared to the use of data from larger geographic units such as census tracts and zip codes. In addition, I could not obtain some retrospective built environment and sociodemographic data, such as traffic and crime for the whole region, which were suggested as important factors by previous researchers (Bowes, 2007, Handy and Clifton, 2001).

3.5 Conclusions

Using the Twin Cities Region as a case study, I examined the relationships between neighborhood type and relative availability of sit-down restaurants and supermarkets, and found a complex and increasingly varied distribution of restaurants and food stores over the past two decades. I observed that the difference in the relative availability of sit-down restaurants between urban core and other neighborhoods are shrinking over time, and I suggest that the fact that the inner city had a higher relative availability of sit-down restaurants than most of the other neighborhoods in the most recent observational year is worth noting. More research using a greater diversity of study regions is needed to identify differences in neighborhood food availability across different settings and contexts. A deeper examination of the co-varying relationship between neighborhood characteristics may help urban planners and neighborhoods improve the likelihood that desired food resources will become more available.

REFERENCES

Bowes, D. R. 2007. A two-stage model of the simultaneous relationship between retail development and crime. *Economic Development Quarterly,* 21**,** 79-90.

Cameron, N., Amrhein, C. G., Smoyer-Tomic, K. E., Raine, K. D. & Chong, L. Y. 2010. Cornering the market: restriction of retail supermarket locations. *Environment and Planning C: Government and Policy,* 28**,** 905-922.

Clary, C. M., Ramos, Y., Shareck, M. & Kestens, Y. 2015. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Preventive medicine,* 71**,** 83-87.

Gustafson, A., Hankins, S. & Jilcott, S. 2012. Measures of the consumer food store environment: a systematic review of the evidence 2000–2011. *Journal of community health,* 37**,** 897-911.

Handy, S. L. & Clifton, K. J. 2001. Local shopping as a strategy for reducing automobile travel. *Transportation,* 28**,** 317-346.

Harvey, H. M. 2017. *Civic Engagement In The Age of Devolution: How Anthropological Approaches Can Help Navigate Grassroots Conflicts.*

Helling, A. & Sawicki, D. S. 2003. Race and residential accessibility to shopping and services. *Housing Policy Debate,* 14**,** 69-101.

James, P., Arcaya, M. C., Parker, D. M., Tucker-Seeley, R. D. & Subramanian, S. 2014. Do minority and poor neighborhoods have higher access to fast-food restaurants in the United States? *Health & place,* 29**,** 10-17.

Jones, M. & Huh, J. 2014. Toward a multidimensional understanding of residential neighborhood: A latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. *Health & place,* 27**,** 134-141.

Lamichhane, A. P., Warren, J., Puett, R., Porter, D. E., Bottai, M., Mayer-Davis, E. J. & Liese, A. D. 2013. Spatial patterning of supermarkets and fast food outlets with respect to neighborhood characteristics. *Health & place,* 23**,** 157-164.

Liese, A. D., Colabianchi, N., Lamichhane, A. P., Barnes, T. L., Hibbert, J. D., Porter, D. E., Nichols, M. D. & Lawson, A. B. 2010. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments. *American journal of epidemiology,* 172**,** 1324-1333.

Lytle, L. A. 2009. Measuring the food environment: state of the science. *American journal of preventive medicine,* 36**,** S134-S144.

Maciag, M. 2015. *Gentrification in America Report* [Online]. Folsom, CA: GOVERNING. Available: http://www.governing.com/gov-data/census/gentrification-in-cities-governing-report.html#citieslist [Accessed Jun 22 2017].

Mercille, G., Richard, L., Gauvin, L., Kestens, Y., Shatenstein, B., Daniel, M. & Payette, H. 2012. Associations between residential food environment and dietary patterns in urban-dwelling older adults: results from the VoisiNuAge study. *Public health nutrition,* 15**,** 2026-2039.

Meyer, K. A., Boone-Heinonen, J., Duffey, K. J., Rodriguez, D. A., Kiefe, C. I., Lewis, C. E. & Gordon-Larsen, P. 2015. Combined measure of neighborhood food and physical activity environments and weight-related outcomes: The CARDIA study. *Health & place,* 33**,** 9-18.

Minneapolis Metropolitan Council 2015. Thrive MSP 2040 Plan: housing policy plan. St. Paul, MN: Metropolitan Counci.

Minnesota Department of Employment and Economic Development. 2014. *State of the Economy* [Online]. St. Paul, MN: Minnesota Department of Employment and Economic Development. Available: https://mn.gov/deed/newscenter/publications/trends/june-2014/state-of-the-economy.jsp [Accessed Jun 22 2017].

Nelson, M. C., Gordon-Larsen, P., Song, Y. & Popkin, B. M. 2006. Built and social environments: associations with adolescent overweight and activity. *American journal of preventive medicine,* 31**,** 109-117.

Powell, L. M., Chaloupka, F. J. & Bao, Y. 2007. The availability of fast-food and full-service restaurants in the United States: associations with neighborhood characteristics. *American journal of preventive medicine,* 33**,** S240-S245.

Powell, L. M., Han, E., Zenk, S. N., Khan, T., Quinn, C. M., Gibbs, K. P., Pugach, O., Barker, D. C., Resnick, E. A. & Myllyluoma, J. 2011. Field validation of secondary commercial data sources on the retail food outlet environment in the US. *Health & place,* 17**,** 1122-1131.

Ramsey, K. & Bell, A. 2014. *Smart location database Version 2.0 User Guide* [Online]. Washington, DC: United States Environmental Protection Agency. Available: https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide [Accessed June 15 2017].

Rummo, P. E., Meyer, K. A., Boone-Heinonen, J., Jacobs Jr, D. R., Kiefe, C. I., Lewis, C. E., Steffen, L. M. & Gordon-Larsen, P. 2015. Neighborhood availability of convenience stores and diet quality: Findings from 20 years of follow-up in the Coronary Artery Risk Development in Young Adults study. *American journal of public health,* 105**,** e65-e73.

U.S. Department of Agriculture. 2017. *Food Environment Atlas* [Online]. Washington, DC USDA's Economic Research Service. [Accessed Jun 22 2017].

Ver Ploeg, M. 2010. *Access to affordable and nutritious food: measuring and understanding food deserts and their consequences: report to Congress*, DIANE Publishing.

Walker, R. E., Keane, C. R. & Burke, J. G. 2010. Disparities and access to healthy food in the United States: a review of food deserts literature. *Health & place,* 16**,** 876-884.

Zenk, S. N. & Powell, L. M. 2008. US secondary schools and food outlets. *Health & place,* 14**,** 336-346.

CONCLUSIONS

The purpose of this research was to examine if and how built environmental factors were related to the use of restaurants and food stores, and to purchase of different healthy food. My goal was to find possible links between built environmental factors and food-related behaviors. A better understanding of the potential bias resulting from the heterogeneity involved in individual food outlet choices may provide the foundation for future longitudinal work. To accomplish the goal, consideration of both neighborhood food availability and the broader built environmental context in which outlets situated are needed.

I observed some evidence that neighborhood food availability was associated with the frequency of use of restaurants and food stores by individual and the amount of food purchased by household. Specifically, I observed a positive association between the availability of neighborhood sit-down restaurants and how often people reported using those restaurants. I found an inverse association between the number of neighborhood convenience stores and expenditures on fruits, which suggests that either these convenience stores did not offer fresh fruits and vegetables or outcompeted outlets that did.

I observed some evidence that the broader built environment context was associated with the frequency of use of restaurants and food stores by individual and the amount of food purchased by household. Specifically, I found an unexpected inverse association between neighborhood street connectivity and how often people reported eating at neighborhood fast food restaurants. I found a positive association between neighborhood street connectivity and expenditures on fruits and vegetables. I also found (paradoxically) that greater regional destination accessibility was associated with greater food expenditures but that the number of supermarkets within a reasonable buffer area was not significantly associated with expenditures on fruits, which suggests that food stores outside of the immediate neighborhood should be investigated in more depth.

I found that although all neighborhoods (except for the urban core) experienced an increase in the percentage of sit-down restaurants between 1993 and 2011, inner city neighborhoods had a greater percentage of neighborhood sit-down restaurants in 2011 than all the other neighborhoods (except for the urban core). I observed that urban and high-income neighborhoods had a smaller percentage of supermarkets than did older suburb neighborhoods in 2001 and 2011.

Significance of this research

I empirically linked some of the rarely examined built environment constructs to the reported use of retail food resources and food purchases under an ecological theory framework. I thus demonstrated that neighborhood context is related to how people use restaurants and food stores and that the food opportunities beyond the home neighborhood are probably also related to food purchases.

Although in some instances, individuals living in a small area with an extremely high numbers of fast food restaurants may eat at those restaurants more often (Forsyth et al., 2012), my results do not support this conclusion. Policies intended to exclude fast food restaurants or to limit the density of fast food restaurants in underserved neighborhoods should therefore be employed with care (McGuire, 2012) as decreasing unhealthy food outlets without introducing alternative food stores could be problematic, particularly for those lacking access to transportation (Sturm and Cohen, 2009, Fox and Horowitz, 2013). My finding concerning convenience stores and fruit purchases should pose a similar caution for policies intended to decrease the relative availability of fast food restaurants, as convenience stores in urban areas may have some additional advantages of spatial accessibility, such as the integration of food shopping into other daily activities such as meeting children and easier parking (Cannuscio et al., 2014).

My work may also indicate the need to discuss of issues related to access to healthy food in inner cities, i.e., will the diet quality of people become worse if they live in inner city neighborhoods with a growing relative number of sit-down restaurants? Will people in inner city neighborhoods in the U.S. act in a fashion similar to the CARDIA participants living in Twin Cities region with respect to eating at sit-down restaurants? If so, there may be cause for concern, as sit-down restaurants do not necessarily provide consistently more healthy meals than fast food restaurants (Saelens et al., 2007). This is not the first time that inner cities in the U.S. have been entangled with food issues. In the early 1970s, the major fast food chains (e.g., McDonald's) realized how profitable their inner city outlets could be and saturated cities with more fast food restaurants, recruiting minority franchisees who would be eligible for federal assistance through minority entrepreneurship and urban renewal programs (Jou, 2017); since that time, fast food restaurants have tended to outnumber grocery stores in American's inner cities (Jou, 2017). Inner cities have been relegated to at best a limited access to supermarkets due to demographic changes in the larger U.S. cities—e.g., affluent households emigrating from inner-cities to suburban areas, thereby decreasing the median income in the inner-cities and forcing nearly one-half of the supermarkets in some of the largest U.S. cities to close (Alwitt and Donley, 1997; Diesenhouse, 1993; Miller, 1994).

My finding concerning sit-down restaurants is in line with the theory of urban vitality that the more restaurants there are, the more attractive the neighborhood for people to engage in neighborhood public activities, which benefits neighborhoods in different ways such as aesthetically pleasing atmosphere and decreased crime rates. Policymakers working on behalf of the inner city regions should therefore evaluate and balance the benefits and costs (e.g., diet quality, street vitality and safety, tax revenue) resulting from the changing food retailing environment. My finding concerning the insignificant association between the presence of neighborhood supermarkets and food purchase does not mean that the lack of supermarket and grocery stores in some inner neighborhoods is not a problem. (Dubowitz et al., 2015) recently reported that diet improved after supermarket introduction in a food desert but not because of supermarket use and suggested that the mechanism behind the improvements in diet is related to the changes in neighborhood satisfaction and perceived access to healthy food.

The fact that both food outlets and the broader built environment were related to food-related behavior suggests the use of growth management tools (e.g., zoning, subdivision) and land use planning to systematically shape the built environment to address issues of access to healthy food. My results suggest the existence of several issues requiring special attention when employing land use tools to improve nutritional health.

The first of these concerns the role played by other built environmental factors (e.g., neighborhood street connectivity) when identifying priority geographic areas for policy intervention. Individuals living in neighborhoods with poor connectivity may use fast food restaurants more often. Failing to consider neighborhood street connectivity may lead to a misestimate of an individual’s exposure to fast food restaurants for those who live in such neighborhoods. Revising the master planning (e.g., how the transportation network should be developed to best serve the future land use pattern), zoning (e.g., facilitating the use of internal roads within the commercial nodes, creating zoning districts where the design of the road within the district can be changed), subdivision regulation (e.g., adding connective roads) and even urban design (Ozbil et al., 2011) to improve road network connectivity may limit residents' exposure to a certain type of neighborhood restaurant. Second, other land use tools such as planned unit developments to encourage more compatible uses (such as grocery stores in residential areas) and to increase the competiveness of grocery stores against convenience stores. My results imply that the presence of convenience stores was as important as grocery stores in influencing the choices people make about healthy food purchases (i.e., fresh fruits or vegetables).

My research has implications for the process of developing food system plans. Although health advocates have proposed that Americans eschew fast food in favor of home cooking, buying fresh and affordable groceries are a challenge that is exacerbated both by transportation access as well as household financial burdens. To address such issue, food system planners may wish to seek to make elements of the industrial food system (e.g., supermarkets, restaurants) more responsive to local communities by looking closely at linkages between the industrial food system and other urban systems (Pothukuchi et al., 2017). If, for example, the diversity of regional accessibility destinations within a reasonable driving distance has the potential to encourage households to purchase more fresh foods, then clearly this is a reason for the development of an urban food system plan beyond the municipal boundary. Greater regional coordination is an example of a strategy that could potentially situate supermarkets along people's daily routes (e.g., commuting corridor) across cities and towns in the region. The focus thus should turn to a comprehensive approach of thinking about food-related behaviors that incorporates all aspects of built environment in the broader area beyond the home neighborhood and an interconnected context not restricted to the food destination (Institute of Medicine, 2012, Raja et al., 2010). This systems approach will require urban planners, transportation experts, public health officials, business owners, and other stakeholders to work together to pursue systematic strategies (both spatial and non-spatial) that regulate land use and the transportation systems related to the spatial movement of consumers of eat-out food and of grocery shoppers, such as retail land use along commuting corridors, the mix of retail and employment in employment neighborhoods.

Limitations

Although this is always a consideration in research, 1 may have omitted important factors that explain residential selection and use of food outlets. Scholars are concerned about the impact of residential self-selection on the use of retail food resources and food purchases (Guthman, 2013). This could have resulted in my misattributing reported use of fast food restaurants to factors other than the possible tautology inherent in the relationship between retail food outlets and their targeted consumer population (Guthman, 2013). The CARDIA questionnaire I used includes an attitudinal question asking if, quote, “grocery stores, restaurants, corner stores” was one of the most important reasons for the respondents to move to the neighborhood or live in the neighborhood for their entire life. This is a crude measure of self-selection partially because the question does not ask explicitly if the participant chose the neighborhood to live due to the presence of a specific type of outlet (e.g., sit-down restaurants).

Also, there may have been the potential for measurement error. In the case of grocery stores in the first paper and food purchase in the second paper, people might forget to report small purchases (for example, beverages or a bag of chips, or an apple) thus underreport use and expenditure, which would have resulted in my potentially underestimating the association between neighborhood availability of grocery stores and the reported use of such stores or food purchase. Another alternative is that residents who were aware that their consumption of fast food restaurants was unhealthy may have underreported their use of fast food restaurants to the CARDIA investigators.

Ecological framework suggests that individual, social environment, physical environment, and macro-level environment all impact eating-related behavior. Although my conceptual framework encompasses variation in individual characteristics (e.g., race, SES), social environment (e.g., neighborhood deprivation factors) and physical environment (i.e., neighborhood food availability and other built environmental factors), the environment components I included in answering such research questions above are by no means exhaustive. For example, I did not include macro-level environment factors such as food marketing and agricultural policies. In characterizing neighborhood food availability, I included measures of only the geographic accessibility of food outlets; I did not consider other dimensions of access or food availability, prices, quality, or marketing inside these outlets, which could also affect food purchasing.

Implications for future research

How people use food resources is highly complex and results from the interplay of multiple influences across different contexts. Some of my results with respect to the built environment, the use of restaurants or food stores, and food purchase are intriguing and point to the need for further more in-depth research to disentangle causal pathways. For example, I did not directly test whether greater street connectivity led to people perceiving certain retail food resources more strongly but suggest that further work be done examining the time duration people are exposed to neighborhood food resources (Scully, 2016) to better understand how people use neighborhood food resources.

Also of interest is an in-depth analysis into the spatial pattern of food outlets, with an assessment of the degree to which fast food restaurants or sit-down restaurants cluster near each other or near homes, schools or highway off-ramps. Just as the number of restaurants is disproportionally distributed across income or urbanicity, restaurants appear to be disproportionally distributed within neighborhoods. Further work examining whether the frequency of use of restaurants differs depending on the spatial distribution of food outlets would extend my understanding of the impact of neighborhood food availability.

Last but not least is the issue of why inner city neighborhoods experienced a greater increase in the percentage of neighborhood sit-down restaurants. Are “more sit-down restaurants” an indicator of gentrification that occurred in some higher-income inner-city neighborhoods? Or could this be due to extensive and relatively evenly-distributed new urban development and the improvement of infrastructure (particularly light rail, the park systems and new sports stadium) inside the Twin Cities (Maciag, 2015) regardless of neighborhood income level? As gentrification is known to increase prices (Zukin et al., 2017), the food at certain establishments could become less affordable for low-income groups in and around gentrified inner city neighborhoods. Further work should thus be done which examines the mechanism underlying the increased presence of sit-down restaurants in inner city neighborhoods, which in turn may facilitate a better understanding of how to improve access to healthy food.

Conclusions

In conclusion, the “big takeaway” from my research is that simply building supermarkets or banning fast food restaurants is not enough; more consideration should be given to other types of food outlets such as sit-down restaurants and convenience stores. A comprehensive plan and collaboration is necessary given that the way that people interact with food environment is complex. Solutions must target everyone to equip all Americans with the resources to make healthy choices.

REFERENCES

Cannuscio, C. C., Hillier, A., Karpyn, A. & Glanz, K. 2014. The social dynamics of healthy food shopping and store choice in an urban environment. *Social Science & Medicine,* 122**,** 13-20.

Dubowitz, T., Ghosh-Dastidar, M., Cohen, D. A., Beckman, R., Steiner, E. D., Hunter, G. P., Flórez, K. R., Huang, C., Vaughan, C. A. & Sloan, J. C. 2015. Diet and perceptions change with supermarket introduction in a food desert, but not because of supermarket use. *Health affairs,* 34**,** 1858-1868.

Forsyth, A., Wall, M., Larson, N., Story, M. & Neumark-Sztainer, D. 2012. Do adolescents who live or go to school near fast-food restaurants eat more frequently from fast-food restaurants? *Health & place,* 18**,** 1261-1269.

Fox, A. M. & Horowitz, C. R. 2013. Best practices in policy approaches to obesity prevention. *Journal of health care for the poor and underserved,* 24**,** 168.

Guthman, J. 2013. Too much food and too little sidewalk? Problematizing the obesogenic environment thesis. *Environment and Planning A,* 45**,** 142-158.

Institute of Medicine 2012. *Accelerating progress in obesity prevention: solving the weight of the nation,* Washington, DC, The National Academies Press.

Jou, C. 2017. *Supersizing Urban America: How Inner Cities Got Fast Food with Government Help*, University of Chicago Press.

Maciag, M. 2015. *Gentrification in America Report* [Online]. Folsom, CA: GOVERNING. Available: http://www.governing.com/gov-data/census/gentrification-in-cities-governing-report.html#citieslist [Accessed Jun 22 2017].

McGuire, S. 2012. Institute of Medicine. 2012. Accelerating progress in obesity prevention: Solving the weight of the nation. Washington, DC: the National Academies Press. Oxford University Press.

Ozbil, A., Peponis, J. & Stone, B. 2011. Understanding the link between street connectivity, land use and pedestrian flows. *Urban Design International,* 16**,** 125-141.

Pothukuchi, K., Arrowsmith, M. & Lyon, N. 2017. Hydraulic Fracturing: A Review of Implications for Food Systems Planning. *Journal of Planning Literature***,** 0885412217733991.

Raja, S., Yin, L., Roemmich, J., Ma, C., Epstein, L., Yadav, P. & Ticoalu, A. B. 2010. Food environment, built environment, and women's BMI: evidence from erie county, New york. *Journal of Planning Education and Research*.

Saelens, B. E., Glanz, K., Sallis, J. F. & Frank, L. D. 2007. Nutrition Environment Measures Study in restaurants (NEMS-R): development and evaluation. *American journal of preventive medicine,* 32**,** 273-281.

Scully, J. Y. 2016. *Human mobility, exposure to the built environment, and health* [Online]. Washington D.C.: University of Washington. Available: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/36862/Scully_washington_0250E_16018.pdf?sequence=1 [Accessed 22 Nov 2017].

Sturm, R. & Cohen, D. A. 2009. Zoning for health? The year-old ban on new fast-food restaurants in South LA. *Health Affairs*, 28**,** w1088-w1097.

APPENDIX 1-1. FOOD OUTLET CLASSIFICATION

Table A1-1. Primary SIC codes from D&B, 8-digit codes shown below.

Food Resource Type	Description	D&B primary SIC code [a]
Fast food Restaurants	Fast food restaurants and stands	58120300
	Box lunch stand	58120301
	Carry-out only (except pizza) restaurant	58120302
	Chili stand	58120303
	Coffee shop	58120304
	Delicatessen (eating places)	58120305
	Drive-in restaurant	58120306
	Fast-food restaurant, chain	58120307
	Fast-food restaurant, independent	58120308
	Food bars	58120309
	Grills (eating places)	58120310
	Hamburger stand	58120311
	Hot dog stand	58120312
	Sandwiches and submarines shop	58120313
	Snack bar	58120314
	Snack shop	58120315
	Pizza restaurants	58120600
	Pizzeria, chain	58120601
	Pizzeria, independent	58120602
Sit-down Restaurants	Ethnic food restaurant	58120100
	American restaurant	58120101
	Cajun restaurant	58120102
	Chinese restaurant	58120103
	French restaurant	58120104
	German restaurant	58120105
	Greek restaurant	58120106
	India/Pakistan restaurant	58120107
	Italian restaurant	58120108
	Japanese restaurant	58120109
	Korean restaurant	58120110
	Lebanese restaurant	58120111
	Spanish restaurant	58120113
	Thai restaurant	58120115
	Vietnamese restaurant	58120116
	Pakistani restaurant	58120117
	Seafood restaurants: include sushi restaurants, oyster bars & seafood shacks	58120114 58120700 58120701 58120702
	Family-owned restaurant	58120500
	Family-owned restaurants, chain	58120501
	Family-owned restaurant, independent	58120502
	Steak house & BBQ restaurants	58120800 58120801 58120802

	Chicken restaurants	58129904
Supermarkets	Supermarkets	54110100
	Supermarkets, chain	54110101
	Supermarkets, independent	54110102
	Supermarkets, greater than 100,000 square feet (hypermarket)	54110103
	Supermarkets, 55,000-65,000 square feet (superstore)	
	Supermarket, 66,000-99,000 square feet	54110104
		54110105
Grocery stores	Grocery store	54110000
	Grocery store, nec	54119900
	Frozen food and freezer plans, except meat	54119903
	Grocery stores, chain	54119904
	Grocery stores, independent	54119905
	Juices, fruit or vegetable	54990202
Convenience	Variety stores	53310000
stores	Convenience stores	54110200
	Convenience stores, chain	54110201
	Convenience stores, independent	54110202
	Gasoline service stations	55410000
	Gasoline service stations, nec	55419900
	Filling stations, gasoline	55419901

Note: [a] D&B has created their own 4-digit extension to the original SIC system as a means to update and expend the system so their customers can more precisely define their business classification.

APPENDIX 1-2. ADDITIONAL INFORMATION ON VARIABLES

Neighborhood food availability

I calculated neighborhood food availability using the numbers of fast-food restaurants, sit-down restaurants, grocery stores, and supermarkets and convenience stores within the 1-km buffer. I generated the counts of food outlets by geocoding the food resources from D&B, a commercial dataset of U.S. business records. I classified fast food restaurants, sit-down restaurants, grocery stores, supermarkets, and convenience stores according to 8-digit SIC codes (Table A1-1 in Appendix 1-1). I calculated the numbers of fast food restaurants, sit-down restaurants, grocery stores, supermarkets and convenience stores within the 3-km buffer using the same method.

Neighborhood street connectivity

I calculated the neighborhood street connectivity by using the link-to-node ratio within the 1-km buffer. The link-to-node ratio is an index of connectivity equal to the number of links divided by the number of nodes. A higher ratio value indicates greater connectivity (Berrigan et al., 2010). I defined links as roadway segments between two nodes, where nodes were intersections or the end of a cul-de-sac. The formula is as follows:

$$\beta = \frac{L}{V}$$

where L is the number of links, V is the number of nodes.

I obtained the road network maps from the ESRI Data and Maps StreetMap North America 2010. Interstate highways and access ramps were eliminated. Non-intersection nodes along contiguous road segments (pseudo-nodes) were also removed. I calculated the link-to-node ratio within the 3-km buffer using the same method.

Population density, neighborhood socioeconomic status (SES) deprivation factor score, vacancy density

I calculated population densities as the total population divided by total land area. I retrieved the total land area and total population for a block group from Census 2000. I interpolated this population density variable from the block group in which participants resided to within my 1-km and 3-km buffers.

I estimated the neighborhood SES deprivation factor score (abbreviated as neighborhood SES deprivation below) using the first factor score from a principal components analysis of the four census indicators of SES (based on a respondent's census tract of residence), which were median household income, % at or below the 150% federal poverty level, % aged 25 or greater with less than a high school education, and % aged 25 or greater with a college degree or higher (Boone-Heinonen et al., 2011b). I obtained the census data of income, poverty, education attainment from the 2005-2009 American Community Survey (ACS) 5-Year Estimates.

I generated density of vacant housing units using the number of vacant housing units (including for rent, rented but not occupied, for sale only, sold but not occupied, for seasonal, recreational, or occasional use, for migrant workers, and other vacant) in the block group divided by the area of block group in which a respondent resided. I obtained the census data of vacant housing units and total land area from the 2005-2009 ACS 5-Year Estimates.

Individual-level sociodemographic variables

I used the information on the highest degree obtained to create a dichotomous indicator of whether the participant received a degree beyond high school. Participants reported their combined family income as falling into one of nine categories (e.g., $5000-11999/year), and I created a measure in U.S. dollars as the midpoint of the selected income category. I defined household size as the number of individuals living in the family. Employment status consisted of two categories, working full-time or part-time and unemployed. I classified participants as married if they reported being currently married or living with someone in a marriage-like relationship. I used the information on reasons for moving to or staying in the participant's current neighborhood to create a dichotomous indicator of whether neighborhood food environment (grocery stores, restaurants, corner stores) was one of the most important reasons for moving to or staying in the neighborhood.

APPENDIX 1-3. REGRESSION RESULTS (3-KM BUFFER)

Table A1-3a Associations between GIS-measured neighborhood fast food availability, neighborhood street connectivity and self-reported frequency of use

GIS-measured exposure	First-step model: perceiving at least one neighborhood fast food restaurant [a] (full Sample)	Second-step model: self-reported frequency of use of neighborhood fast food restaurants [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2007)
Number of fast food restaurants	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Link-to-node ratio	0.64 (0.31, 1.31)	**0.28 (0.12, 0.55)**

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates significant association (P <.05).

Notes: [a] Estimated coefficients of perceiving at least one fast food restaurant in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.

[b] Estimated coefficients of frequency of use of neighborhood fast food restaurants within the participants who perceived at least one fast food restaurant in the participant's neighborhood were adjusted for number of neighborhood sit-down restaurants, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one fast food restaurant in the participant's neighborhood.

Table A1-3b. Associations between GIS-measured neighborhood sit-down restaurant availability, neighborhood street connectivity and self-reported frequency of use

GIS-measured exposure	First-step model: perceiving at least one neighborhood sit-down restaurant [a] (full sample)	Second-step model: self-reported frequency of use of neighborhood sit-down restaurants [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2122)
Number of sit-down restaurants	1.00 (1.00, 1.01)	-0.01 (-0.02, -0.00) [c] -0.00 (-0.01, 0.00) [c] 0.00 (-0.00, 0.01) [c] 0.01 (-0.00, 0.01) [c]
Link-to-node ratio	**3.14 (1.47, 6.73)**	0.71 (0.32, 1.59)

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates significant association (P <.05).

Notes [a] Estimated coefficients of perceiving at least one sit-down restaurant in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.

[b] Estimated coefficients of frequency of use of neighborhood sit-down restaurants within the participants who perceived at least one sit-down restaurant in the participant's neighborhood were adjusted for number of neighborhood fast food restaurants, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one sit-down restaurant in the participant's neighborhood.

[c] The number of sit-down restaurants within the 3-km buffer had four coefficients because the variable violated the proportional odds assumption; the four coefficients indicated the effects of the exposure switching from no use to use yearly, from use yearly to use monthly, from use monthly to use weekly, and from use weekly to use more than once per week, respectively

Table A1-3c. Associations between GIS-measured neighborhood grocery store availability, neighborhood street connectivity and self-reported frequency of use

GIS-measured exposure	First-step model: perceiving at least one neighborhood grocery store (full sample) [a]	Second-step model: self-reported frequency of use of neighborhood grocery stores [b] (restricted sample)
	OR (95% CI) (n=2860)	OR (95% CI) (n=2191)
Number of grocery stores	**0.99 (0.98, 0.99)**	0.99 (0.99, 1.00)
Link-to-node ratio	0.59 (0.26, 1.32)	0.82 (-1.24, 2.87) [b]
		1.05 (-0.40, 2.50) [b]
		0.95 (0.16, 1.75) [b]
		0.35 (-0.36, 1.07) [b]

Abbreviations: OR: odds ratio; CI: confidence intervals. **Bold** indicates significant association (P <.05).

Notes [a] Estimated coefficients of perceiving at least one grocery store in the participant's neighborhood were adjusted for population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the participant's neighborhood.

[b] Estimated coefficients of frequency of use of neighborhood grocery stores within the participants who perceived at least a grocery store in the participant's neighborhood only were adjusted for number of neighborhood supermarkets and convenience stores, population density, neighborhood SES deprivation, vacancy density, family income, race, gender, age, education, employment status, household size, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/ staying in the participant's neighborhood, and the probability of perceiving at least one grocery store in the participant's neighborhood.

APPENDIX 1-4. COLLINEARITY DIAGNOSTICS, BALANCE TEST AND PROPORTIONAL ODDS ASSUMPTION TEST

My collinearity diagnostics results indicated that the variables with high values of variation inflation factor (VIF) were the numbers of outcome food outlet (fast food restaurant, sit-down restaurant, and grocery store) and the interaction term between the outcome food outlet and the link-to-node ratio (Table A1-4a in Appendix 1-4). I used mean-centering (i.e., subtraction of the variable average from the individual data value) to address the multicollinearity issue (Aiken et al., 1991). After mean-centering the numbers of outcome food outlet and the link-to-node ratio, the VIF values of the numbers of outcome food outlet and the interaction terms decreased in the models with the interaction term. The VIF values of the other variables remained almost the same as before the mean-centering. I mean-centered the values of the exposures (number of outcome food outlets and link-to-node ratio) in the final models (*whether the participant perceived at least one neighborhood fast food restaurant/sit-down restaurant/grocery store* model and *frequency of use of neighborhood fast food restaurants/sit-down restaurants/grocery stores* model).

Table A1-4a. Collinearity diagnostics: before and after exposures centered on mean (1-km buffer measures & with interaction term)

	VIF					
	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
Variable	Before [a]	After [b]	Before [a]	After [c]	Before [a]	After [d]
GIS-measured neighborhood food availability measure						
Number of food outlets within 1 km Euclidean buffer around each participant's home						
Fast food restaurants	271.7	10.6	5.5	5.5	---	---
Sit-down restaurants	6.7	6.7	308.4	12.1	---	---
Grocery stores	---	---	---	---	217.5	7.3
Supermarkets and grocery stores	---	---	---	---	2.5	2.5
GIS-measured Neighborhood street connectivity measure						
Link-to-node ratio within 1 km Euclidean buffer around each participant's home	1.7	1.7	1.9	1.9	1.6	1.7
Interaction term between GIS-measured neighborhood food availability and neighborhood street connectivity						
Fast food restaurants*link-to-node ratio	272.5	5.9	---	---	---	---
Sit-down restaurants*link-to-node ratio	---	---	314.0	7.1	---	---
Grocery stores*link-to-node ratio	---	---	---	---	223.1	4.3
Other GIS-measured neighborhood environmental measure						
Population density within 1km Euclidean buffer around each participant's home: 1,000 person/km^2	2.8	2.8	3.1	3.1	5.6	5.6
Neighborhood SES deprivation	1.8	1.8	2.7	2.7	1.9	1.9
Density of vacant housing units: 100 housing units/km^2	2.0	2.0	2.4	2.4	2.0	2.0
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. less than or equal to high school)	1.2	1.2	1.2	1.2	1.2	1.2
Family income, 1,000 $	1.9	1.9	1.9	1.9	1.9	1.9
Black (vs. white)	2.4	2.4	1.4	1.4	2.0	2.0
Female (vs. male)	1.1	1.1	1.1	1.1	1.1	1.1
Household size, person	1.4	1.4	1.4	1.4	1.4	1.4
Age, years	1.1	1.1	1.1	1.1	1.1	1.1
Employed (vs. not employed)	1.1	1.1	1.1	1.1	1.1	1.1
Married (vs. not married)	1.7	1.7	1.7	1.7	1.7	1.7
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood of participants (vs. not one of the most reasons of moving to/staying in the neighborhood)	2.0	2.0	3.4	3.4	2.7	2.7
Probability of perceiving at least one outcome outlet	4.5	4.5	5.8	5.8	4.3	4.3
Mean VIF	34.0	2.9	38.7	3.2	27.8	2.6

Abbreviations: VIF: variation inflation factor. km: kilometer.
Notes [a] All the variables were assigned the original values
[b] Number of fast food restaurants and the link-to-node ratio within the 1-km buffer were centered on mean.
[c] Number of sit-down restaurants and the link-to-node ratio within the 1-km buffer were centered on mean.
[d] Number of grocery stores and the link-to-node ratio within the 1-km buffer were centered on mean.

Table A1-4b. Collinearity diagnostics: before and after exposures centered on mean (1-km buffer measures & without interaction term)

Variable	VIF					
	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
	Before [a]	After [b]	Before [a]	After [c]	Before [a]	After [d]
GIS-measured neighborhood food availability measure						
Number of food outlets within 1 km Euclidean buffer around each participant's home						
Fast food restaurants	5.7	5.7	7.0	7.0	---	---
Sit-down restaurants	6.7	6.7	5.5	5.5	---	---
Grocery stores	---	---	---	---	4.3	4.3
Supermarkets and grocery stores	---	---	---	---	2.5	2.5
GIS-measured Neighborhood street connectivity measure						
Link-to-node ratio within 1 km Euclidean buffer around each participant's home	1.5	1.5	1.8	1.8	1.5	1.5
Other GIS-measured neighborhood environmental measure						
Population density within 1km Euclidean buffer around each participant's home: 1,000 person/km^2	2.8	2.8	3.1	3.1	5.6	5.6
Neighborhood SES deprivation	1.8	1.8	2.7	2.7	1.9	1.9
Density of vacant housing units: 100 housing units/km^2	2.0	2.0	2.3	2.3	1.9	1.9
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. less than or equal to high school)	1.2	1.2	1.2	1.2	1.2	1.2
Family income, 1,000 $	1.9	1.9	1.9	1.9	1.9	1.9
Black (vs. white)	2.4	2.4	1.4	1.4	2.0	2.0
Female (vs. male)	1.1	1.1	1.1	1.1	1.1	1.1
Household size, person	1.4	1.4	1.4	1.4	1.4	1.4
Age, years	1.1	1.1	1.1	1.1	1.1	1.1
Employed (vs. not employed)	1.1	1.1	1.1	1.1	1.1	1.1
Married (vs. not married)	1.7	1.7	1.7	1.7	1.7	1.7
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood of participants (vs. not one of the most reasons of moving to/staying in the neighborhood)	2.0	2.0	3.3	3.3	2.7	2.7
Probability of perceiving at least one outcome outlet	4.3	4.3	5.5	5.5	4.2	4.2
Mean VIF	2.4	2.4	2.6	2.6	2.3	2.3

Abbreviations: VIF: variation inflation factor; km: kilometer.

Notes [a] All the variables were assigned the original values

[b] Number of fast food restaurants and the link-to-node ratio within the 1-km buffer were centered on mean.

[c] Number of sit-down restaurants and the link-to-node ratio within the 1-km buffer were centered on mean.

[d] Number of grocery stores and the link-to-node ratio within the 1-km buffer were centered on mean.

Table A1-4c. Collinearity diagnostics: before and after exposures centered on mean (3-km buffer measures & without interaction term)

Variable	VIF					
	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
	Before [a]	After [b]	Before [a]	After [c]	Before [a]	After [d]
GIS-measured neighborhood food availability measure						
Number of food outlets within 3 km Euclidean buffer around each participant's home						
Fast food restaurants	307.1	19.5	19.4	19.4	---	---
Sit-down restaurants	17.9	17.9	495.7	30.3	---	---
Grocery stores	---	---	---	---	540.7	21.6
Supermarkets and grocery stores	---	---	---	---	6.5	6.5
GIS-measured neighborhood street connectivity measure						
Link-to-node ratio within 3 km Euclidean buffer around each participant's home	2.4	2.4	3.3	3.3	2.2	2.5
Other GIS-measured neighborhood environmental measure						
Population density within 3km Euclidean buffer around each participant's home: 1,000 person/km^2	3.9	3.9	4.0	4.0	10.9	10.9
Neighborhood SES deprivation	1.8	1.8	3.5	3.5	1.9	1.9
Density of vacant housing units: 100 housing units/km^2	1.8	1.8	2.0	2.0	1.7	1.7
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. less than or equal to high school)	1.2	1.2	1.2	1.2	1.2	1.2
Family income, 1,000 $	1.9	1.9	1.9	1.9	2.0	2.0
Black (vs. white)	2.7	2.7	1.4	1.4	2.1	2.1
Female (vs. male)	1.1	1.1	1.1	1.1	1.1	1.1
Household size, person	1.4	1.4	1.4	1.4	1.4	1.4
Age, years	1.1	1.1	1.1	1.1	1.1	1.1
Employed (vs. not employed)	1.1	1.1	1.1	1.1	1.1	1.1
Married (vs. not married)	1.7	1.7	1.7	1.7	1.8	1.8
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood of participants (vs. not one of the most reasons of moving to/staying in the neighborhood)	2.3	2.3	5.0	5.0	3.0	3.0
Probability of perceiving at least one outcome outlet	5.4	5.4	8.9	8.9	5.3	5.3
Mean VIF	41.5	4.3	59.1	5.7	64.7	4.4

Abbreviations: VIF: variation inflation factor; km: kilometer.
Notes [a] All the variables were assigned the original values
[b] Number of fast food restaurants and the link-to-node ratio within the 3-km buffer were centered on mean.
[c] Number of sit-down restaurants and the link-to-node ratio within the 3-km buffer were centered on mean.
[d] Number of grocery stores and the link-to-node ratio within the 3-km buffer were centered on mean.

Balance test

I used only the data of those participants who perceived at least one neighborhood fast food restaurant/sit-down restaurant/grocery store in the *frequency of using neighborhood fast food restaurants/sit-down restaurants/grocery stores* model; those who did not perceive at least one outcome food outlet were excluded from such models. I therefore needed to ensure there is enough overlap (e.g. balance in covariates) between those who perceived at least one outcome food outlet and those who did not. If there was enough overlap between the "perceived" subgroup and the "not perceived" subgroup, the estimated densities of the probability of being perceived versus not perceived would not have too much mass around 0 or around 1 (Busso et al., 2014). I generated the probabilities of perceiving at least one outcome food outlet (i.e., propensity scores) for the two subgroups and plotted them on the same graph. There were enough overlaps between those who perceived at least one neighborhood fast food restaurant/sit-down restaurant/grocery store (Figure A1-4 in Appendix A1-4) and those who did not, suggesting all covariates were largely balanced. Thus, I did not necessarily need to exclude some of the "perceived" participants in the frequency of use models due to significantly different environmental or sociodemographic characteristics. I used kdens command in STATA 14 to draw the plots.

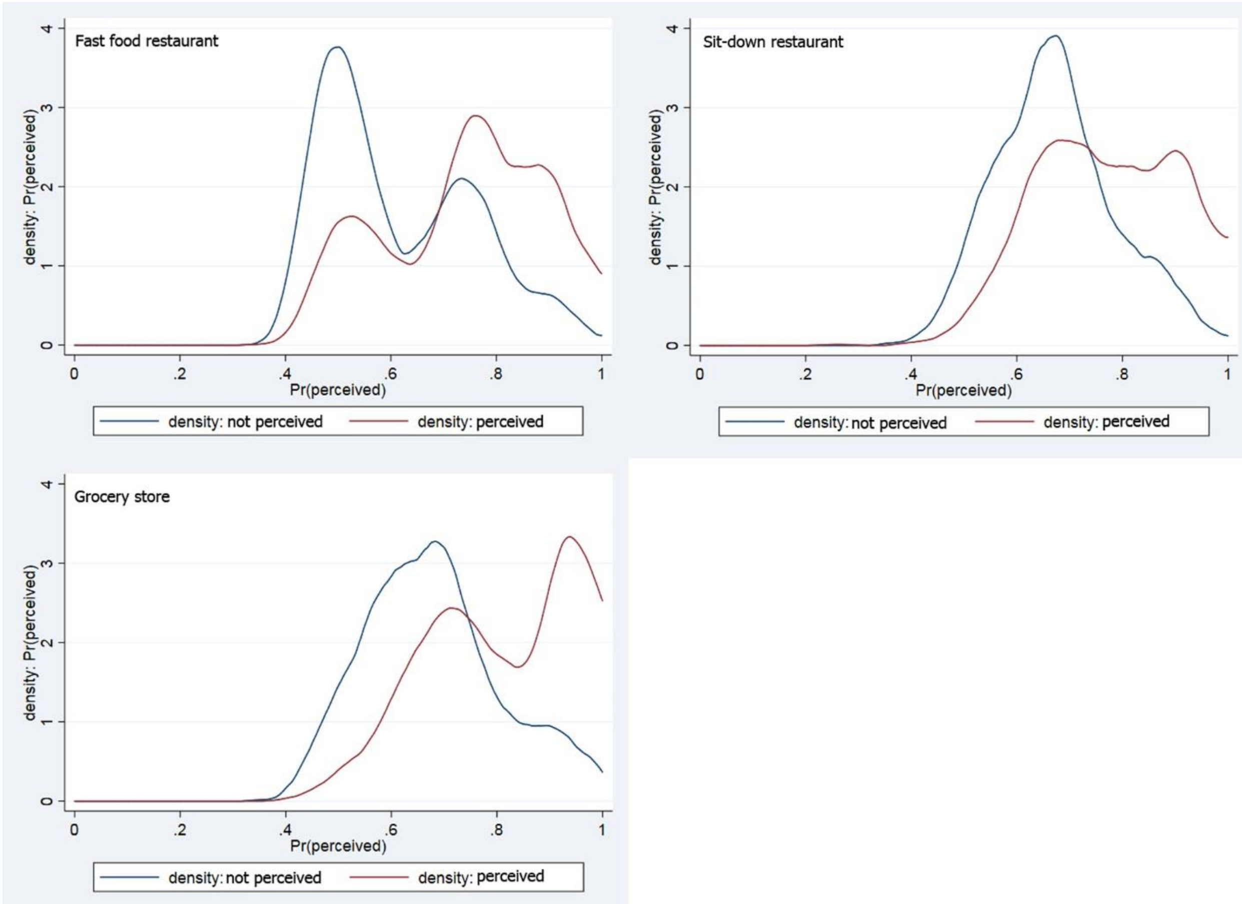

Figure A1-4. Kernel density estimate: probability of perceiving at least one outcome food outlet in the neighborhood

Proportional odds assumption test

Specifically, in the 1-km buffer models (without interaction terms), education attainment and household size in the ordinal frequency of use of neighborhood fast food restaurants model violated the assumption; race, gender, and household size in the ordinal frequency of use of sit-down restaurants model violated the assumption; number of neighborhood grocery stores, number of neighborhood supermarkets and convenience stores, link-to-node ratio, density of vacant housing units, education attainment, race, age, employment status, marital status, and if neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood (abbreviated as *reasons to move* below) in the ordinal frequency of use of grocery stores model violated the assumption (see Table A1-4e in Appendix A1-4).

In the 3-km buffer models, education attainment and household size in the ordinal frequency of use of fast food restaurants model violated the assumption; number of fast food restaurants, number of sit-down restaurants, education attainment, race, gender, household size, and reasons to move in the ordinal frequency of use of sit-down restaurants model violated the assumption; the number of supermarkets and convenience stores, link-to-node ratio, density of vacant housing units, education attainment, family income, gender, and marital status in the ordinal frequency of use of grocery store model violated the assumption (see Table A1-4f in Appendix A1-4).

I therefore relaxed the assumption for the variables that violated the assumption. Each variable that violated the assumption obtained four coefficients which indicated the effects of switching from *no use* to *use yearly*, from *use yearly* to *use monthly*, from *use monthly* to *use weekly*, and from *use weekly* to *use more than once per week*, respectively. I used the gllamm command in Stata 14.0 to test the proportional odds assumption, relax the proportional odds assumption for the variables that violated the assumption, and run the final random intercept mixed effects generalized ordered logistic regressions. I used 8 integration points (nip=8).

I also reported the results of likelihood ratio test of proportional odds assumption with interaction term for the 1-km buffer measures (see Table A1-4d in Appendix A1-4).

Table A1-4d Likelihood ratio test of proportional odds assumption: frequency [a] of use models (1-km buffer measures & with interaction term)

	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
	χ^2	*p*	χ^2	*p*	χ^2	*p*
GIS-measured neighborhood food availability measure						
Number of food outlets within 1 km Euclidean buffer around each participant's home						
Fast food restaurants [b]	0.60	0.896	1.73	0.631	---	---
Sit-down restaurants [b]	0.22	0.974	2.33	0.507	---	---
Grocery stores [b]	---	---	---	---	3.87	0.276
Supermarkets and convenience stores [b]	---	---	---	---	12.93	**0.005**
GIS-measured neighborhood street connectivity measure						
Link-to-node ratio [b] within 1 km Euclidean buffer around each participant's home	1.92	0.589	0.49	0.921	0.61	0.894
Interaction term between GIS-measured neighborhood food availability and neighborhood street connectivity						
Fast food restaurants*link-to-node ratio [b]	0.57	0.903	---	---	---	---
Sit-down restaurants*link-to-node ratio [b]	---	---	3.44	0.329	---	---
Grocery stores*link-to-node ratio [b]	---	---	---	---	4.36	0.225
Other GIS-measured neighborhood environmental measure						
Population density within 1km Euclidean buffer around each participant's home: 1,000 person per square km	4.58	0.205	0.80	0.849	1.52	0.678
Neighborhood SES deprivation	3.55	0.315	3.87	0.276	0.90	0.825
Density of vacant housing units: 100 housing units/km^2	2.54	0.468	0.81	0.847	3.94	0.268
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. lower than or equal to high school)	9.43	**0.024**	8.20	**0.042**	10.83	**0.013**
Family income, 1,000 $	3.80	0.284	1.48	0.687	1.21	0.752
Black (vs. white)	5.30	0.151	8.29	**0.040**	2.87	0.412
Female (vs. male)	0.64	0.888	8.36	**0.039**	6.39	0.094
Household size, person	12.70	**0.005**	9.73	**0.021**	0.94	0.816
Age, years	6.34	0.096	3.06	0.382	12.30	**0.006**
Employed (vs. not employed)	2.97	0.397	1.71	0.634	5.41	0.144
Married (vs. not married)	0.39	0.943	5.81	0.121	16.12	**0.001**
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood (vs. not one of the most reasons of moving to/staying in the neighborhood)	1.61	0.658	7.13	0.068	5.69	0.128
Predicted value of perceiving at least one outcome outlet in the neighborhood	2.19	0.533	2.09	0.555	0.16	0.984

Abbreviations: χ^2: Chi-square; km: kilometer. **Bold** indicates significant association (P <.05).
Notes [a] Frequency of use of outcome (fast food, sit-down or grocery store) outlet is ordered from no use, use yearly, use monthly, use weekly, and use more than once a week.
[b] Numbers of food outlets and link-to-node ratio were centered on mean.

Table A1-4e Likelihood ratio test of proportional odds assumption: frequency [a] of use models (1-km buffer measures without interaction term)

	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
	χ^2	*p*	χ^2	*p*	χ^2	*p*
GIS-measured neighborhood food availability measure						
Number of food outlets within 1 km Euclidean buffer around each participant's home						
Fast food restaurants [b]	0.42	0.936	1.18	0.758	**---**	**---**
Sit-down restaurants [b]	0.20	0.978	2.32	0.508	**---**	**---**
Grocery stores [b]	**---**	**---**	**---**	**---**	31.63	**0.000**
Supermarkets and convenience stores [b]	**---**	**---**	**---**	**---**	44.66	**0.000**
GIS-measured neighborhood street connectivity measure						
Link-to-node ratio [b] within 1 km Euclidean buffer around each participant's home	1.96	0.581	0.49	0.921	14.46	**0.002**
Other GIS-measured neighborhood environmental measure						
Population density within 1km Euclidean buffer around each participant's home: 1,000 person per square km	4.58	0.205	0.61	0.894	3.89	0.273
Neighborhood SES deprivation	3.60	0.308	3.75	0.290	6.65	0.084
Density of vacant housing units: 100 housing units/km^2	2.72	0.437	0.50	0.920	3.94	**0.010**
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. lower than or equal to high school)	9.58	**0.023**	7.51	0.057	25.78	**0.000**
Family income, 1,000 $	3.86	0.277	1.33	0.722	0.56	0.905
Black (vs. white)	5.12	0.163	8.61	**0.035**	13.92	**0.003**
Female (vs. male)	0.64	0.888	8.41	**0.038**	1.01	0.798
Household size, person	12.83	**0.005**	10.37	**0.016**	5.65	0.130
Age, years	6.48	0.091	3.03	0.387	9.55	**0.023**
Employed (vs. not employed)	2.96	0.397	1.74	0.628	7.84	**0.049**
Married (vs. not married)	0.40	0.941	5.85	0.119	40.24	**0.000**
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood (vs. not one of the most reasons of moving to/staying in the neighborhood)	1.54	0.673	6.50	0.090	12.17	**0.007**
Predicted value of perceiving at least one outcome outlet in the neighborhood	2.19	0.533	1.41	0.702	0.02	0.999

Abbreviations: km: kilometer; χ^2: Chi-square. **Bold** indicates significant association (P <.05).
Notes [a] Frequency of use of outcome (fast food, sit-down or grocery store) outlet is ordered from no use, use yearly, use monthly, use weekly, and use more than once a week.
[b] Numbers of food outlets and link-to-node ratio were centered on mean.

Table A1-4f Likelihood ratio test of proportional odds assumption: frequency [a] of use models (3-km buffer measures & without interaction term)

	Fast food restaurant (n=2007)		Sit-down restaurant (n=2122)		Grocery store (n=2191)	
	χ^2	*p*	χ^2	*p*	χ^2	*p*
GIS-measured neighborhood food availability measure						
Number of food outlets within 3 km Euclidean buffer around each participant's home						
Fast food restaurants [b]	2.04	0.564	12.15	**0.007**	---	---
Sit-down restaurants [b]	2.53	0.469	14.50	**0.002**	---	---
Grocery stores [b]	---	---	---	---	3.45	0.328
Supermarkets and convenience stores [b]	---	---	---	---	10.99	**0.012**
GIS-measured neighborhood street connectivity measure						
Link-to-node ratio [b] within 3 km Euclidean buffer around each participant's home	1.11	0.774	2.03	0.567	17.71	**0.001**
Other GIS-measured neighborhood environmental measure						
Population density within 3km Euclidean buffer around each participant's home: 1,000 person per square km	2.74	0.433	1.29	0.732	1.78	0.620
Neighborhood SES deprivation	3.57	0.312	5.46	0.141	3.01	0.390
Density of vacant housing units: 100 housing units/km^2	1.94	0.584	2.09	0.555	8.15	**0.043**
Measures of self-reported individual-level sociodemographic and reasons to moving to/staying in the neighborhood						
Education greater than high school (vs. less than or equal to high school)	9.41	**0.024**	7.93	**0.048**	9.19	**0.027**
Family income, 1,000 $	3.50	0.320	1.09	0.781	9.54	**0.023**
Black (vs. white)	5.55	0.136	8.66	**0.034**	1.70	0.637
Female (vs. male)	0.82	0.844	8.78	**0.032**	9.04	**0.029**
Household size, person	11.70	**0.009**	9.38	**0.025**	0.76	0.859
Age, years	6.29	0.099	3.83	0.281	6.90	0.075
Employed (vs. not employed)	2.27	0.518	1.72	0.632	4.87	0.182
Married (vs. not married)	0.68	0.878	4.78	0.188	19.80	**0.000**
Neighborhood food environment (grocery stores, restaurants, corner stores) is one of the most important reasons of moving to/staying in the neighborhood (vs. not one of the most reasons of moving to/staying in the neighborhood)	0.65	0.884	8.89	**0.031**	6.79	0.079
Predicted value of perceiving at least one outcome outlet in the neighborhood	2.07	0.558	4.71	0.194	0.14	0.987

Abbreviations: χ^2: Chi-square; km: kilometer. **Bold** indicates significant association (P <.05).

Notes [a] Frequency of use of outcome (fast food, sit-down or grocery store) outlet is ordered from no use, use yearly, use monthly, use weekly, and use more than once a week.

[b] Numbers of food outlets and link-to-node ratio were centered on mean.

APPENDIX 2-1. SAMPLE

Nielsen had two types of households who reported the purchase of fresh fruits (abbreviated as *fruits* below) and fresh vegetables (abbreviates as *vegetables* below), magnet panelists and non-magnet panelist. Magnet panelist reported products that do or do not use standard UPC codes whereas non-magnet panelist reported products that use standard UPC codes only. Because many random-weight fresh fruits and vegetables do not use standard UPC codes, I only use the expenditure information from n=27,422 magnet households. I excluded 4,559 for not residing in an MSA. I excluded 131 due to missing covariate information. Among the remaining 22,732 Nielsen households (which I describe below as the *raw study sample*), I further excluded 284 due to extremely low or high values for purchases of both (fruits and vegetables). The annual purchase of fruits and vegetables by household is a high frequency item compared to low frequency products such as clothes and furniture. However, I still observed that 1,229 and 917 households spent less than $10 on fruits and vegetables in my raw study sample, respectively. In addition, 34 and 24 households spent more than $ 1,000 on fruits and vegetables in the raw study sample, respectively. To avoid undue influence by some outlying points on the regression, I excluded the households with extremely large or small expenditures on both fruits and vegetables (both simultaneous). That is to say, the households with extremely large or small expenditures on fruits but not on vegetables, or vice versa, were still included. I defined extreme values of expenditure on fruits or vegetables as those below the 2^{st} percentile or above the 98^{th} percentiles. My final sample size was n=22,448. I did not use below 1^{st} percentile or above the 99^{th} percentiles as the cut off points because I will still have some households with zero purchase of fruits or vegetables if I used the 1^{st} percentile as the cut-off point, which did not facilitate the log transformation of expenditures on fruits and vegetables.

APPENDIX 2-2. MEASURES

Purchase of fresh fruits and vegetables

The purchase of fresh fruits and vegetables using standard UPC codes were categorized as *fresh produce* by Nielsen (Table A2-2a). By choosing the products under the category of fresh produce, I explicitly excluded fruits and vegetables that were dried, tinned, bottled, frozen, or refrigerated, which were under other department categories such as dry grocery, frozen food, and deli.

The purchase of fresh fruits and vegetables using non-standard UPC codes were categorized as *magnet* by Nielsen (Table A2-2a). Under the products of magnet data, I used brand_desr= reference card fruits and brand_desc= reference card vegetables to identify fruits and vegetables, respectively. Other non-standard UPC (i.e., magnet) products included baked goods, prepared foods, cheese, meat/poultry/fish, coffee, flora, and etc., according to the description of products (variable name= brand_desr) in the product file. Items in the categories "reference card fruits" and "reference card vegetables" were not described in detail, i.e., whether the fruits or vegetables were fresh, dried, tinned, frozen, or refrigerated and etc. As magnet data was known for random weighted (loose) products (Allcott et al., 2017, Einav et al., 2008, Oster, 2015, Zhen et al., 2009), I assumed that I could ignore the proportion of random weighted magnet dried, frozen fruits or vegetables. I therefore specified all the magnet products with brand_descr=reference card fruits and brand_descr=reference card vegetables as fresh fruits and fresh vegetables, respectively.

Table A2-2a. Fresh produce type and Nielsen production group module description

Food type	Department [a]	Product module/brand [a]
Fruits using standard UPC codes	Fresh produce	fresh apples, fresh cranberries, fresh grapefruits, fresh kiwi, fresh oranges, fresh strawberries, fresh tomatoes, and fresh fruits-remaining.
Fruits using non-standard UPC codes	Magnet	Magnet data with brand_descr=reference card fruits
Vegetables using standard UPC codes	Fresh produce	fresh carrots, fresh cauliflower, fresh celery, fresh lettuce, fresh garlic, fresh mushrooms, fresh onions, fresh potatoes, fresh radishes, fresh spinach, fresh sprout, and fresh vegetables-remaining.
Vegetables using non-standard UPC codes	Magnet	Magnet data with brand_descr=reference card vegetables

Note: [a] Department, product module, and brand are all Nielsen defined product codes.

In calculating the annual expenditure on fruits or vegetables by household, I linked the product file to the purchase file using the UPC numbers (variable name 1=upc, variable name 2=upc_ver_uc) as the joint identifying numbers to create a purchase-product file. Upc_ver_uc indicated different versions of upc. I then linked the purchase-product file to the trip file using the trip number (variable name=trip_code_uc) as the joint identifying number to create a trip-purchase-product file. I linked the trip-purchase-product file to the household sociodemographic file using the household number (variable name=household_code) as the joint identifying number to create a household-trip-purchase-product file.

Availability of neighborhood supermarkets and convenience stores

I classified the supermarkets and convenience stores according to primary six-digit standard industrial classification (SIC) codes (Table A2-2b).

Table A2-2b. Primary 6-digit SIC codes from ReferenceUSA used in the analysis for year 2010

Food Resource Type	Description	ReferenceUSA primary SIC code
Supermarkets	Supermarkets	541101
Convenience stores	Variety stores	533100
	Snack products	541102
	Convenience stores	541103
	Gasoline service stations	554100
	Gas station and convenience stores	554199
Fast food restaurants	Fast food restaurants and stands	581203
	Pizza restaurant	581206
Sit-down Restaurants	Fine dining	581201
	Family restaurants	581205
	Seafood restaurants	581207
	Steak and barbecue restaurant	581208
Schools	Schools	821103
Child care services	Child care services	835101
Churches	Churches	866107

Regional destination accessibility

I retrieved the regional destination accessibility from the Smart Location Database (SLD). The SLD measured regional destination accessibility as the number of jobs within a 45-minute automobile travel time (network travel time decay weighted) for each block group. The SLD used the employment information in the InfoUSA 2011 and the street network information in the NAVSTREETS to generate measures of job opportunities in each reachable block group and the travel time between each origin block group and all the destination block groups using network analysis models. SLD then generated the value of regional destination accessibility by decaying the employment at destinations by the distance decay curve and summed for each origin block group (Ramsey and Bell, 2014). I spatially linked the residence zipcode to the SLD block group using ArcGIS 10.5. I was thus able to link the block-group-level regional destination accessibility calculated by the SLD to the corresponding household.

Neighborhood street connectivity

I retrieved the neighborhood street connectivity calculated by the SLD. The SLD used the road network information in the NAVSTREETS to measure neighborhood street connectivity as the total weighted number of street intersections divided by total land area in the block group. The formula used by the SLD to calculate the weighted street connectivity in the block group is as follows:

$$street\ connectivity = d1 * 0.667 + d2 + d3 * 0.667 + d4$$

where *d1* is the number of multi-modal intersections having three legs per square mile, *d2* is the number of multi-modal intersections having four or more legs per square mile, *d3* is the number of pedestrian-oriented intersections having three legs per square mile, and *d4* is the number of pedestrian-oriented intersections having four or more legs per square mile. To reflect the connectivity for pedestrian and bicycle travel, the SLD assigned a weight of zero to auto-oriented intersections to reflect the fact that they are a barrier to pedestrian and bicycle mobility (Ramsey and Bell, 2014). Similarly, the SLD assigned lower weights to three-way intersections to reflect the fact that they do not promote street connectivity as effectively as four way intersections (Ramsey and Bell, 2014). I interpolated this street connectivity variable from the block group in which household resided (the centroid of residential zip code tabulation area) to within the 5-km buffer and the 3-km buffer, separately.

Household-level covariates

Households reported the highest degree obtained by the female head of household as falling into one of seven categories, which I combined to generate a new education variable with three categories, which were 1(high school or below), 2 (college or higher), and 3 (no female head). I chose the female head of household rather than the male head of household because maternal education previously has been shown to be an important determinant of child diet (Zhen et al., 2009, Crawford et al., 1995, Northstone and Emmett, 2005, Hendricks et al., 2006). Households reported their household income as falling into one of 20 categories, which I combined to generate a new income variable with three categories, which were less than $20,000, $20,000-$49,999, and $50,000 or more. Households reported the race identity of their households as falling into one of four categories, including white, black, Asian and other. Households reported the household size as falling into one of nine categories; I combined these nine categories to generate a new household variable with four categories, which were 1 (one member), 2 (two members), 3 (three members), and 4 (four members or above). Households reported the marital status of household head as falling into four categories, which were married, widowed, divorced/separated, and single.

Data analyses

Of the final study sample (obs=22,448), my final sample size was n=21,824 and n=21,824 in the fruit and vegetable models, respectively. I ran each model twice, once for availability of neighborhood supermarkets and convenience stores, regional destination accessibility, availability of neighborhood destinations, neighborhood destination diversity, and neighborhood street connectivity within the 5-km buffer and again for the same factors within the 3-km buffer. Fruit and vegetable models included several household-, neighborhood- and area-level covariates (see the method part in the main text). See Table 2.2, Table 2.3, Table A2-3a, and Table A2-3b in Appendix 2-3 for the regression results within the 5-km buffer and the 3-km buffer, respectively.

In my sensitivity analysis, of the raw study sample (obs=22,732), I further excluded 451 households due to extremely low or high values for purchases of fruits and vegetables (both simultaneously). See Appendix 2-1 for the definition of raw study sample. I defined an extreme value for expenditure on fruits or vegetables as those below the 3rd percentile or above the 97th

percentile. The final study sample in the sensitivity analysis is 22,281 (22,732-451=22,281) households. My final sample size was n=21,370 and n=21,368 in the fruit and vegetable models, respectively. In my sensitivity analysis, I ran each model twice, each time of the same factors as in my main analysis, but (again, similar to my main analysis) once for the factors within the 5-km buffer and again for the same factors within the 3-km buffer. Fruit and vegetable models included several household-, neighborhood- and area-level covariates (see the method part in the main text). See Table A2-3c and A2-3d in Appendix 2-3 for the regression results for the 5-km buffer.

References

Aiken, L. S., West, S. G. & Reno, R. R. 1991. *Multiple regression: Testing and interpreting interactions*, Sage.

Allcott, H., Diamond, R. & Dubé, J.-P. 2017. The geography of poverty and nutrition: Food deserts and food choices across the United States. National Bureau of Economic Research.

Berrigan, D., Pickle, L. W. & Dill, J. 2010. Associations between street connectivity and active transportation. *International journal of health geographics,* 9**,** 20.

Boone-Heinonen, J., Roux, A. V. D., Kiefe, C. I., Lewis, C. E., Guilkey, D. K. & Gordon-Larsen, P. 2011b. Neighborhood socioeconomic status predictors of physical activity through young to middle adulthood: the CARDIA study. *Social science & medicine,* 72**,** 641-649.

Busso, M., DiNardo, J. & McCrary, J. 2014. New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics,* 96**,** 885-897.

Crawford, P. B., Obarzanek, E., Schreiber, G. B., Barrier, P., Goldman, S., Frederick, M. M. & Sabry, Z. 1995. The effects of race, household income, and parental education on nutrient intakes of 9-and 10-year-old girls NHLBI growth and health study. *Annals of epidemiology,* 5**,** 360-368.

Einav, L., Leibtag, E. & Nevo, A. 2008. *On the accuracy of Nielsen Homescan data*, USDA, Economic Research Service.

Hendricks, K., Briefel, R., Novak, T. & Ziegler, P. 2006. Maternal and child characteristics associated with infant and toddler feeding practices. *Journal of the American Dietetic Association,* 106**,** 135-148.

Northstone, K. & Emmett, P. 2005. Multivariate analysis of diet in children at four and seven years of age and associations with socio-demographic characteristics. *European journal of clinical nutrition,* 59**,** 751.

Oster, E. 2015. Diabetes and Diet: Purchasing Behavior in Response to Health News.

Ramsey, K. & Bell, A. 2014. *Smart location database Version 2.0 User Guide* [Online]. Washington, DC: United States Environmental Protection Agency. Available: https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide [Accessed June 15 2017].

Zhen, C., Taylor, J. L., Muth, M. K. & Leibtag, E. 2009. Understanding differences in self-reported expenditures between household scanner data and diary survey data: a comparison of Homescan and consumer expenditure survey. *Review of Agricultural Economics,* 31**,** 470-492.

APPENDIX 2-3. COEFFICIENTS OF ASSOCIATION BETWEEN BUILT ENVIRONMENT CHARACTERISTICS, HOUSEHOLD-, NEIGHBORHOOD- AND AREA-LEVEL, AND FRUITS AND VEGETABLES PURCHASED BY NIELSEN HOUSEHOLD

Table A2-3a. Coefficients of cross-sectional associations between built environment characteristics (3-km buffer), household-, neighborhood- and area-level, and fruits purchased by Nielsen household (obs=22,448 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 3-km buffer				
0 (ref)	---	---	---	---
1	0.05 (0.02)	**0.015**	0.02 (0.02)	0.232
2+	0.01 (0.03)	0.780	-0.02 (0.03)	0.390
Broader built environmental context				
Availability of convenience stores, 10 counts, 3-km buffer			-0.06 (0.02)	**0.001**
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.00 (0.00)	**0.000**
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 3-km buffer			0.00 (0.00)	0.334
Neighborhood destination diversity: Entropy, 10 percent, 3-km buffer			0.00 (0.00)	0.157
Neighborhood street connectivity: 10 intersections per square mile, 3-km buffer			0.01 (0.00)	**0.005**
Urbanicity				
Urbanized area (ref)			---	---
Urban cluster			-0.08 (0.03)	**0.022**
Non-urban area			-0.00 (0.02)	0.988

Abbreviation: SE: standard error. **Bold** indicates significant association (P <.05)

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 2nd percentile or greater than the 98th percentile. The 21, 824 households in the fruit model were not necessarily the same 21,824 households in the vegetable model; there were 624 households in the fruit model but not in the vegetable model, and vice versa.

[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

Table A2-3b. Coefficients of cross-sectional associations between built environment characteristics (3-km buffer), household-, neighborhood- and area-level, and vegetables purchased by Nielsen household (obs=22,448 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 3-km buffer				
0 (ref)	---	---	---	---
1	0.04 (0.02)	**0.048**	0.02 (0.02)	0.224
2+	0.02 (0.02)	0.366	0.00 (0.03)	0.985
Broader built environmental context				
Availability of convenience stores, 10 counts, 3-km buffer			-0.02 (0.02)	0.193
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.00 (0.00)	0.434
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 3-km buffer			0.00 (0.00)	0.603
Neighborhood destination diversity: Entropy, 10 percent, 3-km buffer			0.00 (0.00)	0.603
Neighborhood street connectivity: 10 intersections per square mile, 3-km buffer			0.01 (0.00)	**0.007**
Urbanicity				
Urbanized area (ref)			---	---
Urban cluster			-0.01 (0.03)	0.705
Non-urban area			0.00 (0.01)	0.987

Abbreviation: SE: standard error. **Bold** indicates significant association (P <.05)

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 2nd percentile or greater than the 98th percentile. The 21, 824 households in the fruit model were not necessarily the same 21,824 households in the vegetable model; there were 624 households in the fruit model but not in the vegetable model, and vice versa.

[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

Table A2-3c. Coefficients of cross-sectional associations between built environment characteristics (5-km buffer), household-, neighborhood- and area-level covariates, and fruits purchased by Nielsen household (obs=22,281 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 5-km buffer				
0 (ref)	---	---	---	---
1	0.04 (0.02)	**0.020**	0.02 (0.02)	0.220
2+	0.07 (0.02)	**0.000**	0.03 (0.02)	0.183
Broader built environmental context				
Availability of convenience stores, 10 counts, 5-km buffer			-0.02 (0.01)	**0.006**
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.00 (0.00)	**0.001**
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 5-km buffer			0.00 (0.00)	0.658
Neighborhood destination diversity: Entropy, 10 percent, 5-km buffer			-0.00 (0.00)	0.991
Neighborhood street connectivity: 10 intersections per square mile, 5-km buffer			0.01 (0.00)	**0.006**
Urbanicity				
Urbanized area (ref)			---	---
Urban cluster			-0.08 (0.03)	**0.016**
Non-urban area			-0.01 (0.01)	0.642

Abbreviation: SE: standard error. **Bold** indicates significant association (P <.05)
Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.
Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 3rd percentile or greater than the 97th percentile.
[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

Table A2-3d. Coefficients of cross-sectional associations between built environment characteristics (5-km buffer), household-, neighborhood- and area-level covariates, and vegetables purchased by Nielsen household (obs=22,281 [a])

Characteristics	Presence of neighborhood supermarkets only [b]		Presence of neighborhood supermarkets and broader built environmental context [b]	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Presence of supermarkets, 5-km buffer				
0 (ref)	---	---	---	---
1	0.03 (0.02)	**0.043**	0.03 (0.02)	0.082
2+	0.04 (0.02)	**0.016**	0.03 (0.02)	0.205
Broader built environmental context				
Availability of convenience stores, 10 counts, 5-km buffer			-0.01 (0.01)	0.204
Regional destination accessibility: Jobs within 45-min automobile travel time , 10,000 jobs			0.00 (0.00)	0.082
Availability of neighborhood destinations: Total fast food restaurants, sit-down restaurants, schools, child care services, and churches, 10 counts, 5-km buffer			-0.00 (0.00)	0.585
Neighborhood destination diversity: Entropy, 10 percent, 5-km buffer			-0.00 (0.00)	**0.026**
Neighborhood street connectivity: 10 intersections per square mile, 5-km buffer			0.01 (0.00)	**0.029**
Urbanicity				
Urbanized area (ref)			---	---
Urban cluster			-0.01 (0.03)	0.618
Non-urban area			-0.00 (0.01)	0.978

Abbreviation: SE: standard error. **Bold** indicates significant association (P <.05)
Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.
Notes: [a] I excluded who reported extremely low or high values for purchases of fruits and vegetables (both simultaneously), defined here as less than the 3rd percentile or greater than the 97th percentile.
[b] Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one children in the family, number of employed household members (household head excluded).

APPENDIX 2-4. BALANCE TEST

Balance test

I used only the data of those magnet households who reported non-standard UPC products, which included random weighted (loose) items such as fruits, vegetables, meats, and in-store baked goods; those who did not report non-standard UPC products were excluded from such models. I therefore needed to ensure there is enough overlap (e.g. balance in covariates) between magnet and non-magnet households. If there was enough overlap between the magnet households and the non-magnet households, the estimated densities of the probability of being magnet household versus non-magnet household would not have too much mass around 0 or around 1 (Busso et al., 2014). I generated the probabilities of being magnet (i.e., propensity scores) for the two subgroups and plotted them on the same graph. There were enough overlaps between those who were magnet households and those who were not (Figure A2-3 in Appendix 2-4), suggesting all covariates were largely balanced. Thus, I did not necessarily need to exclude some of the magnet households in the expenditure on fruits or vegetables models due to significantly different household-level sociodemographic characteristics. I used kdens command in STATA 14 to draw the plots.

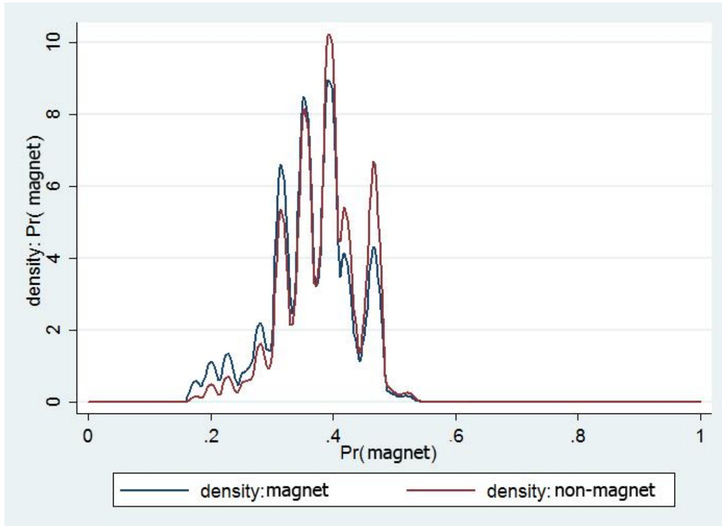

Figure A2-4. Kernel density estimate: probability of being magnet households

References

Busso, M., DiNardo, J. & McCrary, J. 2014. New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics,* 96**,** 885-897.

APPENDIX 3-1. SPECIFIC NEIGHBORHOOD CHARACTERISTICS [A] BY NEIGHBORHOOD TYPE CLASSIFIED IN 1993

Table A3-1. Specific neighborhood characteristics [a] by neighborhood type classified in 1993

	Urban core (n=63)	Inner city (n=103)	Urban (n=335)	Aging suburb (n=579)	High-income suburb (n=331)	Suburban edge (n=672)
Built environment						
Residential population density, 1,000 person/km^2	**5.26±1.79 (2.79)**	3.68±1.24 (1.62)	2.45±1.03 (0.70)	1.66±0.80 (0.11)	0.87±0.48 (-0.48)	**0.54±0.46 (-0.72)**
Employment population density, 1,000 person/km^2	**3.32±1.37 (3.34)**	1.37±0.62 (0.74)	1.41±0.57 (0.81)	0.84±0.39 (0.05)	0.49±0.28 (-0.42)	**0.31±0.27 (-0.66)**
Mix of land use, % [b]	55.19±22.15 (0.39)	48.41±23.91 (0.15)	**55.61±24.84 (0.41)**	55.54±25.39 (0.41)	38.13±28.62 (-0.22)	**30.08±24.78 (-0.51)**
% single-family housing [c]	**20.98± 21.61 (-1.19)**	40.29±24.53 (-0.62)	33.46±28.08 (-0.82)	50.63±30.98 (-0.31)	75.71±28.08 (0.43)	**84.18±21.62 (0.68)**
Sociodemographic						
% population aged under 14 [d]	**7.49±3.85 (-2.29)** [a]	28.09±7.51 (0.91)	17.16±4.51 (-0.79)	20.08±4.18 (-0.34)	20.27±3.42 (-0.31)	**28.13±2.98 (0.91)**
% population aged between 15 and 29 [d]	**50.76±14.21 (3.69)**	25.11±4.82 (0.19)	26.84±7.02 (0.42)	22.64±3.88 (-0.15)	**19.32±4.03 (-0.60)**	22.56±3.61 (-0.16)
% population aged between 30 and 44 [d]	**22.41±7.10 (-1.00)**	23.63±3.31 (-0.75)	29.28±4.22 (0.45)	23.69±2.75 (-0.73)	25.63±3.11 (-0.32)	**30.83±3.40 (0.78)**
% population aged between 45 and 64 [d]	**10.13±4.85 (-1.36)**	12.44±2.83 (-0.90)	15.01±3.18 (-0.38)	17.93±3.42 (0.20)	**24.48±3.49 (1.51)**	14.60±3.08 (-0.46)
% population aged 65 or above [d]	9.25±6.30(-0.10)	10.71±4.46 (0.12)	11.66±4.61 (0.27)	**15.61±6.18 (0.87)**	10.30±4.96 (0.06)	**3.87±2.31 (-0.92)**
Median household income, $ 1,000 [e]	20.01±6.11 (-1.48)	**1.76±5.59 (-1.68)**	33.83±8.06 (-0.38)	31.04±5.93 (-0.60)	**52.13±11.98 (1.08)**	45.79±6.54 (0.57)
% white race [d]	83.29±9.50 (-0.64)	**44.21±16.98 (-3.65)**	89.90±7.91 (-0.13)	92.95±6.60 (0.10)	96.75±1.85 (0.39)	**96.89±2.13 (0.40)**
% black race [d]	8.87±7.25 (0.53)	**35.05±20.11 (3.39)**	5.47±5.51 (0.16)	2.53±3.19 (-0.16)	**0.89±1.08 (-0.34)**	0.90±1.18 (-0.34)
% population with a college education or above [d]	**70.81±10.80 (0.87)**	**38.55±9.19 (-1.25)**	66.80±12.24 (0.60)	45.06±10.29 (-0.82)	70.18±10.42 (0.83)	59.28±12.12(-0.11)

Bold indicates the highest/lowest value of z score in the six types of neighborhoods.
Notes. [a] Mean ± standard error (mean z-score) of neighborhood characteristics measured at the census block group level.
[b] The mix of land use was measured by 3-tier land use entropy (denominator set to the static 3 land use types in the census block group), which used three land use categories (residential, employment and retail) to calculate mix of land use in the census block group.
[c] Percent of single-family housing relative to total single-family and multi-family housings.
[d] The denominators of percent of population aged under 14, aged between 15 and 29, aged between 30 and 44, aged between 45 and 64, aged 65 or above, population with a college education or above, white race and black race were total population in the census block group.
[e] The median household income in 1993 and 2001 were adjusted for inflation to compare with that in 2011.

APPENDIX 3-2. COEFFICIENTS OF THE MODELS OF PERCENTAGE OF SIT-DOWN RESTAURANTS AND PERCENTAGE OF SUPERMARKETS

Table A3-2a. Predicted multivariable-adjusted model coefficients of associations among the percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants, neighborhood type in 1993, interaction of the latter with time elapsed, and time elapsed from years 1993, 2001 and 2011: Twin Cities Region of Minnesota

Predictors	**b (95% CI)**	**P value**
Neighborhood type in 1993 [a]		
Urban core	23.02 (13.18, 32.85)	**0.000**
Inner city (Ref)	…	…
Urban	0.23 (-6.80, 7.27)	0.948
Aging suburb	-0.73 (-7.66, 6.19)	0.835
High-income suburb	-1.91 (-9.48, 5.65)	0.620
Suburban edge	-4.94 (-12.28, 2.41)	0.188
Neighborhood type in 1993: time elapsed [b]		
Urban core	-1.13 (-1.83, -0.44)	**0.001**
Inner city (Ref)	…	…
Urban	-0.55 (-1.04, -0.07)	**0.026**
Aging suburb	-0.60 (-1.07, -0.14)	**0.010**
High-income suburb	-0.37 (-0.86, 0.12)	0.137
Suburban edge	-0.41 (-0.86, 0.05)	0.081
Time elapsed [c]	0.82 (0.39, 1.25)	**0.000**
Covariates		
Change in residential population density, 1,000 person/km^2	0.96 (-0.11, 2.02)	0.078
Change in income, 1,000 US dollar	0.07 (-0.03, 0.17)	0.164
Change in percent of white	0.05 (-0.04, 0.14)	0.282
Change in percent of single family housing	0.02 (-0.03, 0.07)	0.480
Total of sit-down restaurants and fast food restaurants, count	2.99 (2.70, 3.28)	**0.000**
Constant	14.50 (7.21, 21.78)	**0.000**

Abbreviations: b: model effect; CI: confidential interval. **Bold** indicates significant association (P <.05). N=6,249.

Notes [a] The coefficient of neighborhood type in 1993 shows if other types of neighborhoods had a greater percentage of sit-down restaurants than the reference neighborhood type (inner city) in 1993.

[b] Time elapsed in 1993, 2001, and 2011 is defined as 0, 8, and 18, respectively. The coefficient of the interaction term between neighborhood type in 1993 and the time elapsed shows if other types of neighborhoods experienced a greater increase in the percentage of sit-down restaurants than the reference neighborhood type (inner city).

[c] The coefficient of time elapsed refers to the effect of time on the reference neighborhood type (inner city). The coefficient of time elapsed shows if the reference neighborhood type experienced a significant change in the percentage of sit-down restaurants between 1993 and 2011.

Table A3-2b. Predicted multivariable-adjusted model coefficients of associations among the percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores, neighborhood type in 1993, interaction of the latter with time elapsed, and time elapsed from years 1993, 2001 and 2011: Twin Cities Region of Minnesota

Predictors	**b (95% CI)**	**P value**
Neighborhood type in 1993 [a]		
Urban core	1.87 (-2.50, 6.24)	0.401
Inner city (Ref)	…	…
Urban	0.46 (-2.67, 3.59)	0.773
Aging suburb	1.69 (-1.39, 4.77)	0.282
High-income suburb	1.30 (-2.07, 4.66)	0.450
Suburban edge	0.19 (-3.08, 3.46)	0.909
Neighborhood type in 1993: time elapsed [b]		
Urban core	0.02 (-0.29, 0.33)	0.910
Inner city (Ref)	…	…
Urban	-0.07 (-0.29, 0.15)	0.546
Aging suburb	0.00 (-0.21, 0.21)	0.990
High-income suburb	-0.15 (-0.37, 0.07)	0.182
Suburban edge	-0.00 (-0.21, 0.20)	0.988
Time elapsed [c]	**0.24 (0.05, 0.43)**	**0.014**
Covariates		
Change in residential population density, 1,000 person/km^2	0.04 (-0.43, 0.51)	0.869
Change in income, 1,000 US dollar	-0.04 (-0.08, 0.01)	0.101
Change in percent of white	0.02 (-0.02, 0.06)	0.356
Change in percent of single-family housing	**-0.03 (-0.05, -0.01)**	**0.016**
Total of supermarkets, grocery stores and convenience stores, count	**1.83 (1.51, 2.16)**	**0.000**
Constant	-0.83 (-4.08, 2.42)	0.616

Abbreviations: b: model effect; CI: confidential interval. **Bold** indicates significant association ($P < .05$). N=6,249.
Notes [a] The coefficient of neighborhood type in 1993 shows if other types of neighborhoods had a greater percentage of supermarkets than the reference neighborhood type (inner city) in 1993.
[b] Time elapsed in 1993, 2001, and 2011 is defined as 0, 8, and 18, respectively. The coefficient of the interaction term between neighborhood type in 1993 and the time elapsed shows if other types of neighborhoods experienced a greater increase in the percentage of supermarkets than the reference neighborhood type (inner city).
[c] The coefficient of time elapsed refers to the effect of time on the reference neighborhood type (inner city). The coefficient of time elapsed shows if the reference neighborhood type experienced a significant change in the percentage of supermarkets between 1993 and 2011.

APPENDIX 3-3. P VALUES FOR THE CHANGES OF DIFFERENCES IN ESTIMATED MEAN OF PERCENTAGE OF SIT-DOWN RESTAURANTS/SUPERMARKETS FOR EACH NEIGHBORHOOD TYPE PAIR BETWEEN TWO OBSERVATION YEARS

Table A3-3. P values for the changes of difference in estimated mean [a] of percentage of sit-down restaurants/supermarkets for each neighborhood type pair between two observational years

	P value [b]					
	Sit-down restaurant			Supermarket		
	1993 vs. 2001	2001 vs. 2011	1993 vs. 2011	1993 vs. 2001	2001 vs. 2011	1993 vs. 2011
Urban core vs. inner city	**0.00** [c]	**0.00**	**0.00**	0.91	0.91	0.91
Urban core vs. urban	0.06	0.06	0.06	0.53	0.53	0.53
Urban core vs. aging suburb	0.07	0.07	0.07	0.90	0.90	0.90
Urban core vs. high-income suburb	**0.01**	**0.01**	**0.01**	0.22	0.22	0.22
Urban core vs. suburban edge	**0.01**	**0.01**	**0.01**	0.88	0.88	0.88
Inner city vs. urban	**0.03**	**0.03**	**0.03**	0.55	0.55	0.55
Inner city vs. aging suburb	**0.01**	**0.01**	**0.01**	0.99	0.99	0.99
Inner city vs. high-income suburb	0.14	0.14	0.14	0.18	0.18	0.18
Inner city vs. suburban edge	0.08	0.08	0.08	0.99	0.99	0.99
Urban vs. aging suburb	0.74	0.74	0.74	0.31	0.31	0.31
Urban vs. high-income suburb	0.29	0.29	0.29	0.30	0.30	0.30
Urban vs. suburban edge	0.33	0.33	0.33	0.33	0.33	0.33
Aging suburb vs. high-income suburb	0.13	0.13	0.13	**0.03**	**0.03**	**0.03**
Aging suburb vs. suburban edge	0.13	0.13	0.13	0.96	0.96	0.96
High-income suburb vs. suburban edge	0.80	0.80	0.80	**0.03**	**0.03**	**0.03**

Bold indicates significant association ($P < .05$).

Notes [a] Multivariable mixed effects regression modeling percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants/percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores in each neighborhood as a function of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white and percent of single-family housing since 1993, total sit-down and fast food restaurants/total supermarkets, grocery stores, and convenience stores, and a random intercept for each neighborhood.

[b] P value for two-tailed Student's t-test of difference in difference of estimated mean of percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants/percentage of supermarkets relative to total supermarkets, grocery stores and convenience stores

[c] The p value of 0.00 indicates that the difference of estimated mean of percentage of sit-down restaurants relative to total sit-down restaurants and fast food restaurants between urban core and inner city in 1993 is significantly different from that in 2001.